

PSYCHOLOGICAL REVIEW PUBLICATIONS

Psychological Review

EDITED BY

HERBERT S. LANGFELD, PRINCETON UNIVERSITY

S. W. FERNBERGER, UNIV. OF PENNSYLVANIA (*J. of Exper. Psychol.*)

W. S. HUNTER, CLARE UNIVERSITY (*Psychol. Index*)

JOSEPH PETERSON, GEO. PEABODY COLLEGE (*Psychol. Monog.*)

J. A. MCGEOCH, UNIV. OF MISSOURI (*Psychol. Bull.*)

CONTENTS

The Mechanism of the Assembly of Behavior Segments in Novel Combinations Suitable for Problem Solution: CLARK L. HULL, 219.

Scientific Measurement and Psychology: DOUGLAS MCGREGOR, 246.

On the Tactile Perception of Vibration Frequencies: WALTHER JOEL, 267.

The Experimental Approach to Concept Learning: KENNETH L. SMOKE, 274.

A Reply to 'Sign-Gestalt or Conditioned Reflex?': NEAL E. MILLER, 280.

The Two-Factor Theory and its Criticism: HENRY E. GARRETT, 293.

PUBLISHED BI-MONTHLY

FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION

BY THE PSYCHOLOGICAL REVIEW COMPANY

PRINCE AND LEMON STS., LANCASTER, PA.

AND PRINCETON, N. J.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

PUBLICATIONS OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*J. Exper. Psych.*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index and Abstracts*)
HENRY T. MOORE, SKIDMORE COLLEGE (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Review*)
JOSEPH PETERSON, GEORGE PEABODY COLLEGE (*Monographs*)
JOHN A. MCGEOCH, UNIVERSITY OF MISSOURI (*Bulletin*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

is a comprehensive bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: \$5.50 (Foreign, \$5.75). Index: \$4.00 per volume.
Journal: \$7.00 (Foreign, \$7.25). Monographs: \$6.00 per volume (Foreign, \$6.30).
Bulletin: \$6.00 (Foreign, \$6.25). Abstracts: \$6.00 (Foreign, \$6.25).
Abnormal and Social: \$5.00 (Foreign, \$5.25). Single copies \$1.50.
Current numbers: Journal, \$1.25; Review, \$1.00; Abstracts, 75c; Bulletin, 60c.

COMBINATION RATES

Review and Bulletin: \$10.00 (Foreign, \$10.50).
Review and J. Exp.: \$11.00 (Foreign, \$11.50).
Bulletin and J. Exp.: \$12.00 (Foreign, \$12.50).
Review, Bulletin, and J. Exp.: \$16.00 (Foreign, \$16.75).
Review, Bulletin, J. Exp., and Index: \$19.00 (Foreign, \$19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, NEW JERSEY

THE PSYCHOLOGICAL REVIEW

THE MECHANISM OF THE ASSEMBLY OF BEHAVIOR SEGMENTS IN NOVEL COMBINATIONS SUITABLE FOR PROBLEM SOLUTION¹

BY CLARK L. HULL

Institute of Human Relations, Yale University

I. THE GENERAL PROBLEM OF ADAPTIVE NOVELTY IN MAMMALIAN BEHAVIOR

Many persons have been puzzled by the paradox of the presumptive fertility and originality of the processes of reasoning on the one hand, as contrasted with the remarkable sterility of the syllogism on the other. It has been urged that if one already knows a major premise such as,

All men are mortal,

and that if an organism passing by the name of Socrates manifests traits generally characteristic of a man, it requires no particular originality or perspicacity to conclude that Socrates will himself prove ultimately to be mortal also. Fertility, originality, invention, insight, the spontaneous use of implements or tools—these things, clearly, do not lie in the syllogism. The fact that it has been found possible to construct a relatively simple mechanism of sliding disk segments of sheet metal which will solve automatically, *i.e.*, exhibit the conclusions logically flowing from, all of the known syllogisms and which will automatically detect all of the formal fallacies,²

¹ The writer is indebted to the members of his seminar for a number of valuable criticisms and suggestions, notably, to Dr. T. L. McCulloch, Mr. S. D. S. Spragg, and Dr. J. B. Wolfe. Dr. N. R. F. Maier also read a preliminary draft of the manuscript.

² Such a mechanism has been designed and constructed by the author, but a description has not yet been published.

emphasizes the crudely mechanical characteristics of the syllogism.³ The solution of the paradox is, of course, that the genuinely creative and novelty-producing portions of the reasoning processes take place in advance of the emergence of the substance or materials of the syllogism; *i.e.*, the solution consists in the *assembly*, from the considerable store of such materials presumably possessed by the more versatile and adaptive organisms, of the particular set of premises which are relevant to the problem situation in question. An understanding of the dynamics of the presumably numerous forms of intelligence,⁴ insight, thought, and reasoning must therefore be sought not in the mechanism and use of the syllogism as such, but in this period *antecedent* to the explicit emergence of the material which may be susceptible later of being arranged in the form of a syllogism.⁵ It accordingly becomes our task to discover the principles by which, on the occasion of need,

³ So far as the writer is able to see, there is no *a priori* impossibility of constructing a mechanism which will display genuine thinking capacity. Indeed, it is expected with some confidence that such mechanisms will ultimately be constructed. But when, and if, this takes place the thinking mechanism will surely be of a far more subtle and complex character than a mere logic machine consisting of sliding disks (1, 2, 11, 12).

⁴ By a curious hemianopsia, workers in intelligence testing have confined their activities almost exclusively to the field of tests, and their methods have been largely limited to the computation and manipulation of correlation coefficients. The test-correlation approach is admirably adapted for immediate practical application; but, even when supplemented by the more powerful methods of Spearman (24), Kelley (9), Hotelling (3), and Thurstone (28), it is difficult to see how such a method can possibly yield deductions (and thus, explanations) concerning such fundamental phenomena as those propounded by Maier's analysis. Present indications are that the task of developing a systematic scientific theory of intelligence in this sense will be performed mainly by students of animal behavior, a group almost entirely distinct from that professionally engaged in intelligence testing. It may be that it would not be economical for many individuals to attempt to pursue both, though a more intimate linking of the two fields would seem to be desirable for the robust development even of intelligence testing, as is illustrated by the work of E. L. Thorndike (27). Lippmann, in his attack on psychologists engaged in intelligence testing (13), was probably wrong when he urged that they could not measure a thing of whose nature they had little comprehension. Nevertheless it would seem that improvements in the validity of such measurements might be facilitated by a serious effort to understand the numerous mechanisms which presumably lie behind the various manifestations of intelligence.

⁵ As a matter of fact, John Stuart Mill long ago pointed out that the syllogism is primarily a device for testing the accuracy of reasoning processes which have already taken place (20, Book II).

there emerge the habit segments or premises in the particular combination necessary for problem solution.

Something of the theoretical urgency of this problem is brought home to us when we consider the blind chaos which would result if, in problem situations, premises or habit segments should appear by pairs as if drawn by chance from a huge urn from the supply possessed by the organism. The probability that a particular pair of numbers would be drawn from a supply of 100 on any particular occasion is something like 1 in 10,000. Chance is evidently an element in intelligent behavior but, clearly, the dice of chance must be loaded in some way or problems would never get solved. In short, any adequate theory of higher adaptive behavior must show how the dice are loaded, *i.e.*, how the characteristics of the problem situation are able to evoke the particular combination of acts which alone will serve to extricate the organism from its difficulty.

II. THE SPECIFIC PROBLEM OF THE ADAPTIVE ASSEMBLY OF HABIT SEGMENTS IN NOVEL COMBINATIONS

As a rule it is more economical and generally effective to attack difficult problems by investigating at the outset their simpler aspects and manifestations rather than their more complex ones. In accordance with this principle, we shall begin our analysis of the dynamics of novelty in mammalian behavior by the consideration of a concrete form of intelligent or insightful behavior to which attention has been directed by Norman Maier. In this connection he has remarked significantly, "... the combination of two patterns in the solution of a problem is at the bottom of theories of reasoning that make reasoning more than 'trial-and-error' . . ." (17). As a result of ingenious experimental procedures, Dr. Maier believes that he has demonstrated the existence of such capacity in the albino rat.

In order to make the substance of Maier's fundamental bit of analysis explicit, let us consider a somewhat modified and conventionalized version of his experimental arrangement

(17). This may be understood with the aid of Fig. 1. Sections *R*, *U*, *X*, and *H* represent enclosed boxes each of distinct shape. Distinctive cutaneous stimuli are provided for the animals' feet by the character of the floor of each box, on the assumption that a characteristic stimulus will lead to a distinct

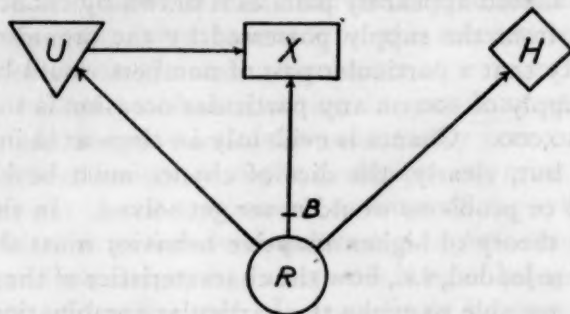
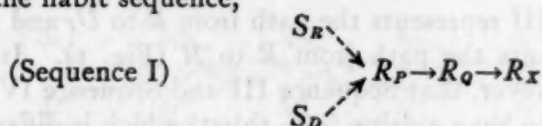


FIG. 1. Diagrammatic representation of a series of locomotor paths representing a somewhat conventionalized form of Maier's 'reasoning' experiment with rats (17). The pathways are enclosed, as are the areas *R*, *X*, *U*, and *H*. The floor of *X* would be of soft, flossy silk, say; that of *U*, of cold metal with rough, sharp points; that of *H*, of polished warm metal; and that of *R*, of several layers of thin rubber dam. When training on one path is taking place, all others are closed.

reaction while in each box, and so favor the arousal of partially distinct anticipatory reactions (5). For the same reason, additional special devices calculated to induce markedly characteristic and distinct postures in the animals whenever traversing the respective boxes, should probably be provided (21). The animal will be trained first to go from *R* to *X*, and from *U* to *X* as distinct habits, and will receive food. Next he will be trained to go from *R* to *U* and from *R* to *H* as distinct habits for the reward of water, say. Following this there should probably be given a few more forced runs on *UX* and *RX* with food reward, after which the animals should be given an opportunity to choose between *RU* and *RH* a number of times to determine objectively the relative strength of the tendency of each individual to enter the respective alleys. Then, with all the four paths open but with a barrier in *RX* at point *B* (Fig. 1), the rat will be placed at *R*, very hungry but with the thirst drive thoroughly satiated. Insight in the sense

here used will be shown by the animals' tending, in general, to choose path RU (and then UX) rather than path RH a significantly greater proportion of the trials than was the case on the control choices.

Let us now, for convenience and definiteness in exposition, represent schematically our conception of the essentials of the organization of the several habit segments generated by the foregoing procedure. We shall begin with that leading from R to X , Fig. 1. We shall suppose that a mammalian organism in an external stimulus situation, S_R , and with an internal stimulus situation (drive), S_D , possesses at the end of some training the habit sequence,



Here S_R represents the visual, auditory, and other external stimuli coming to the organism from box R , and S_D represents proprioceptive stimuli such as result from hunger cramps of the digestive tract, say. R_X represents the final or consummatory reaction which abolishes S_D and thus naturally terminates the cycle. In this sense R_X may be said to be the solution of the problem jointly presented by S_R and S_D . R_P and R_Q represent the locomotor and other activity involved in traversing path $R \rightarrow X$.

Let us suppose, also, that in a different external stimulus situation, S_U , but one possessing the same drive as that of Sequence I, i.e., S_D , there has been formed a second behavior sequence which terminates in the same solution or goal reaction as Sequence I:

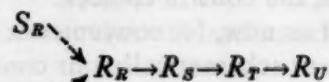


This corresponds to the habit segment leading from U to X , Fig. 1. Here S_U represents the visual, olfactory, and other external stimuli coming to the organism from box U (Fig. 1). R_U , R_V , and R_W represent the activity of traversing the path

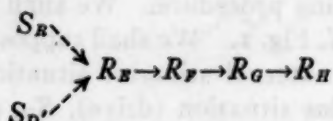
from box U to box X , and R_X represents the final or consummatory response as in Sequence I.

In addition there are postulated two further sequences, both originating at S_R :

(Sequence III)

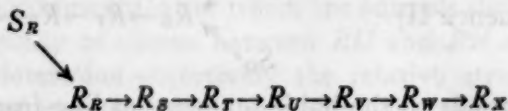


(Sequence IV)



Sequence III represents the path from R to U , and Sequence IV represents the path from R to H (Fig. 1). It is to be noted, however, that Sequence III and Sequence IV are each supposed to have a drive (e.g., thirst) which is different from that of Sequences I and II. We accordingly represent the drive of Sequences III and IV by S_D . R_U and R_H represent that portion of the final or consummatory segments of the respective sequences which are characteristic and distinctive of each.

Now, suppose that Sequence I should be prevented from taking place by the barrier placed between S_R and R_P as indicated in Fig. 2 and point B , Fig. 1. How can the solution (reaction R_X) to the problem thus precipitated be brought about? In external stimulus situation S_R the organism now has but two choices, Sequence III or Sequence IV. If Sequence IV is taken, this act leads only to R_H , which for the drive, S_D , is a mere blind alley and a failure, since it does not eliminate S_D . But if Sequence III is taken, this leads to Sequence II which, in turn, leads to R_X , the solution, thus:



In any realistic consideration of this problem it is important to note that there is always a possibility that the organism will find its way from S_R to S_U by mere chance (trial-and-

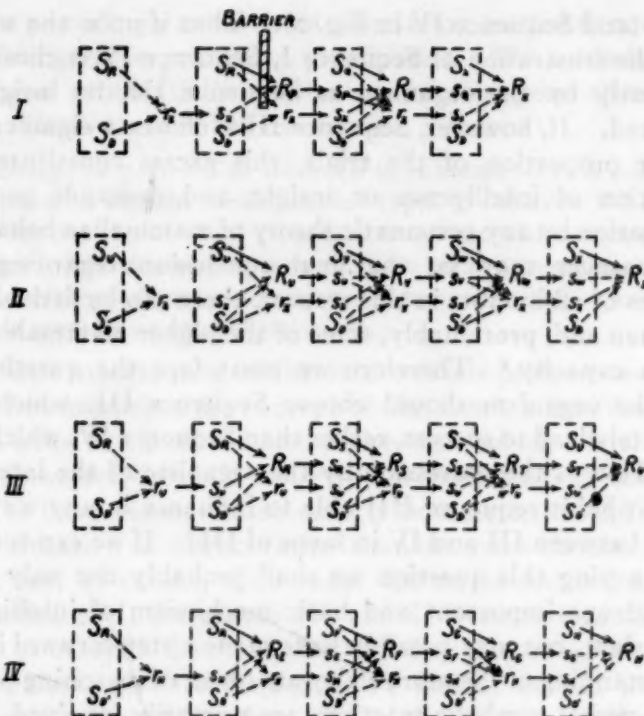


FIG. 2. Diagrammatic representation of the major immediate excitatory tendencies in action sequences I, II, III, and IV, which correspond to Paths $R \rightarrow X$, $U \rightarrow X$, $R \rightarrow U$, and $R \rightarrow H$, respectively, of Fig. 1. Remote excitatory tendencies are not represented. The barrier is assumed to have been placed before R_P after Sequence I had become fairly well practiced and was working smoothly.

error), in which case Sequence II would presumably follow at once and the problem would be solved for the organism but no insight would be involved. Since the trial-and-error type of behavior is almost certain to be present to a greater or less degree in situations such as the one here supposed, it is necessary to have in the experimental set-up some unambiguous opportunity for trial-and-error to operate relatively uncomplicated by other factors in order to have a basis for comparison and possible contrast with a situation in which insight may be an appreciable additional element. It is to be noted that such an opportunity for the manifestation of pure trial-and-error behavior is provided for by the path from R to H of

Fig. 1 (and Sequence IV in Fig. 2). Thus if upon the whole, after the frustration of Sequence I, Sequence IV is chosen as frequently by the organism as Sequence III, no insight is indicated. If, however, Sequence III is chosen a significantly greater proportion of the trials, this excess constitutes an indication of intelligence or insight and demands genuine explanation by any systematic theory of mammalian behavior.

Whatever may be the final conclusion regarding the abilities of albino rats in this respect, there can be little doubt that man and, presumably, some of the higher mammals show such a capacity.⁶ Therefore we must face the question of why the organism should choose Sequence III, which will ultimately lead to success, rather than Sequence IV, which will not. How is the possession by the organism of the latent or implicit habit sequence (II) able to influence in any way the choice between III and IV in favor of III? If we can succeed in answering this question we shall probably not only have isolated one important and basic mechanism of intelligence and insight, but may possibly have made a step forward in the understanding of the more elaborate forms of reasoning proper where verbal symbolic reactions are primarily involved.

III. MAIER'S *Gestalt* INTERPRETATION

Before proceeding to the consideration of our own interpretation of this fascinating problem, it will be well to glance briefly at that offered by Maier himself. He states that when he attempted to account for the animals' alleged preference for the path from *R* to *U* over that from *R* to *H*, by means of associative principles, he encountered difficulties. He argues, first, that since *S_R* has associations leading both to Sequences III and IV (*R*→*U* and *R*→*H*, Fig. 1), it should be expected that the animals would choose the one as readily as the other. But insightful behavior demands that some preference should be shown to III over IV. Thus he concludes that association cannot account for the phenomenon (17, p. 92).

⁶ Recently Wolfe and Spragg (29) have reported results which, taken in conjunction with certain characteristics of Maier's technique, lead them to question whether the albino rat can combine habit segments, as supposed by Maier.

He accordingly turns from association to *Gestalt* concepts as offering more promising possibilities. His remarks concerning a substantially similar situation appear also to be applicable to the one before us:

The concept of patterns or *Gestalten* thus seems to be a necessary assumption to explain these complex types of behavior. The fact that a rat can choose the . . . means to an end without previously having reached this end by any of these means, seems to make a pattern concept almost a necessity. A temporal chain is not sufficient; it must be an immediate whole.

A little examination of the nature of our problem seems to show, however, that the concept of *Gestalten* leads to the same sterile issue as the naïve associative view which Maier quite properly rejected. It is to be observed that Maier does not claim to have demonstrated exactly how patterns lead in the above situation to correct rather than incorrect choices. One might say, of course, that under the stress created by the frustration of I, Sequences III and II become fused into a *Gestalt* or unity in the sense that together they somehow solve the problem, whereas Sequences IV and II do not so fuse. But such an interpretation of *Gestalten* as an explanatory principle would naïvely beg the question, since the deduction of this fusion from more basic principles is the essence of the problem before us. Such a utilization of the concept of *Gestalten* would be a mere tautological gesture; it would merely re-assert the fact of problem solution in a new terminology without in any sense deducing the outcome from any principles whatever. It is true that the history of science reveals many cases of such naïve procedures, though, so far as the writer is aware, no one has ever put forward this particular argument.

But if ordinary association and *Gestalt* principles have both failed, what other possibility remains? The writer is inclined to the view that the principles of association between stimuli and responses, particularly as revealed in modern conditioned-reaction experiments,⁷ offer a possibility of explana-

⁷ In order to correct a frequent misunderstanding, due presumably to the wide dissemination of the views of J. B. Watson, the writer wishes to make it quite clear

tion in a manner which Maier's analysis failed to take into consideration. We shall now proceed to an examination of this possibility.

IV. A SUGGESTED STIMULUS-RESPONSE EXPLANATION OF THE ADAPTIVE ASSEMBLY OF HABIT SEGMENTS

A somewhat detailed stimulus-response analysis of action Sequences I, II, III, and IV, as sketched above, is given in Fig. 2. The S 's at the top of each diagram represent typical stimuli as received from the external receptors. S_D and $S_{D'}$ represent persisting internal stimuli such as hunger, thirst, sex, etc., and the s 's represent proprioceptive stimulations. Broken arrows indicate acquired or learned tendencies, whereas solid lines indicate what are presumed to be innate or unacquired tendencies.

In this connection it is to be noted that an anticipatory goal reaction appears in the first segment of each of these sequences (5). For example, in the first sequence there appears r_x , which is supposed to be a relatively inconspicuous component of the goal reaction R_x , brought forward to the beginning of the series presumably through its association with S_D or through the action of trace reactions while in their early stages (or both), and which, once there, becomes associated with S_R . In the same way we find r_x at the beginning of Sequence II (having originally been a component of R_x),⁸ r_U at

that neither here nor in any previous publications has he assumed that the more complex forms of behavior are synthesized from reflexes which play the rôle of building blocks. This may or may not be true. His working hypothesis is, rather, that the *principles of action* discovered in conditioned reaction experiments are also operative in the higher behavioral processes. The compound adjective in the expression, 'conditioned-reflex principles,' accordingly refers to the locus of *discovery* of the principles rather than to their locus of *operation*.

⁸That terminal or goal and near-goal reactions do come forward in behavior sequences is amply substantiated by a number of investigations, among which may be mentioned those by Lumley (14, 15, 16), Mitchell (22), Spragg (25), and Miller (21). In some of the investigations just mentioned, the anticipatory (or antedating) reactions actually come forward in their entirety and supplant reactions properly belonging in the positions in question. Spragg's study also shows anticipatory tendencies not only for goal reactions, but for sub-goal reactions as well. Moreover, it suggests that sub-goal reactions tend to come forward with less vigor than goal reactions, quite as the goal gradient hypothesis would lead us to expect.

the beginning of Sequence III (having originally been a component of R_U), and r_H at the beginning of Sequence IV (having originally been a component of R_H). It is assumed, further, that similar components of all reactions in any given series tend to come forward and become associated with the external stimulus component in the same manner, though because of the difficulty of representing them in detail only the anticipatory goal reactions are shown in Fig. 2. A complete set of anticipatory reactions, emanating from both goal and sub-goal reactions, are, however, shown in the first segment of Fig. 3.

Fig. 3 is designed to show the dynamics of the situation at point S_R after the barrier comes in to frustrate and extinguish Sequence I. This is regarded as the crucial point of the theoretical problem. A study of Fig. 3, in conjunction with Fig. 2 from which it was derived, shows that the ordinary principles of the association of stimuli and responses, when applied in a thorough-going fashion, actually lead in a logical and straightforward manner to the expectation that an organism capable of functional anticipatory reactions (e.g., r_U) would be more likely to react with the sequence

$$R_R \rightarrow R_S \rightarrow R_T \rightarrow R_U \rightarrow R_V \rightarrow R_W \rightarrow R_X$$

and consequently with a successful solution, than by the reaction

$$R_E \rightarrow R_F \rightarrow R_G \rightarrow R_H$$

which is not a solution.

At bottom the decisive factor in this competition (Fig. 3) is the presence of s_U , with its excitatory tendency to evoke R_R rather than R_E . But s_U , in turn, was necessarily dependent upon the presence of r_U , an anticipatory or antedating reaction. The details of the dominance of r_U over the other anticipatory reactions, and of R_R over R_E , are explained in the legend of Fig. 3, which must be traced through in detail if the reader is to understand the deduction. It may be noted that r_U is a pure-stimulus or symbolic act (4), since the only function that it performs is to release by its action the propriocep-

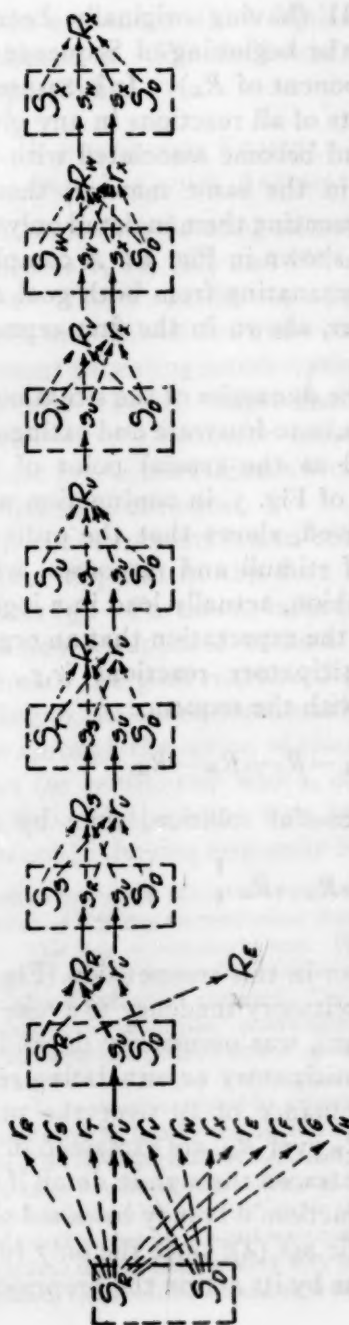


FIG. 3. Diagrammatic representation of the dynamics of the situation at the problem point S_g after the frustration and extinction of the excitatory tendencies originally emanating from it to the sequence $R_p \rightarrow R_q \rightarrow R_x$. Note that all of the excitatory tendencies here represented are taken from Fig. 2. Thus the excitatory tendencies from S_g to r_R, r_S, r_T , and r_U come from Sequence III, those from S_g to r_E, r_F, r_G , and r_H come from Sequence IV, and those from S_D to r_U, r_V, r_W , and r_X come from Sequence II. Referring, now, to the above figure, it may be seen that of the several anticipatory reactions just mentioned, r_U , only, has more than one excitatory tendency. The one from S_g comes from the beginning of Sequence III, whereas the one originating in S_D comes from the first reaction of Sequence II ($S_D \rightarrow R_U$). All the other r 's are accordingly assumed to play minor rôles and are dismissed from further consideration.

At the next segment there appears another critical situation, the actual competition between R_g and R_g . It may be seen from the diagram that R_g has two excitatory tendencies, both originally functioning at the beginning of Sequence III, whereas R_g , its competitor, has only one excitatory tendency, that originating from the beginning of Sequence IV. Because of the advantage of the bond emanating from S_U to R_g , the latter reaction must have the advantage over R_g . Thus the presence of S_U raises the reaction from the status of trial-and-error into the realm of insight or intelligence without the intervention of any special psychic agent.

tive stimulus s_U , which ultimately leads to an intelligent rather than to a stupid reaction.⁹

Here, then, we appear to have laid bare before us the mechanism of *one* type of intelligence or insight. It turns out, in fact, to be associative in nature, though distinctly not in the simple and direct fashion conceived by Maier. Thus there has apparently evolved on a purely physical basis a type of reaction which has sometimes been supposed possible only by a kind of miraculous intervention of some non-physical (psychic) agent called mind or consciousness. To state the same thing in other words, we appear to have before us here a deduction of insight in terms such that it might conceivably be constructed by a clever engineer as a non-living—even an inorganic—mechanism.

V. SEVEN COROLLARIES

A study of Figs. 3 and 4 leads to the statement of a number of corollaries.

(1) One of these takes its origin from the remoteness and general tenuity or feebleness of the mechanism leading to the resolution of the competition in favor, first, of r_U and, second, of R_R , upon which this supposed act of intelligence depends (Fig. 3). Because of the obviously unstable nature of this mechanism, it might naturally be expected that it would be easily upset, that even weak antagonistic trial-and-error impulses would frequently over-ride it, and that, as a consequence, only too often it would fail completely. Thus if the excitatory tendency from S_R to R_R chanced to be much stronger than that from S_R to R_U , the 'intelligence' mechanism of s_U to R_R would be over-ridden. The relative rarity of genuinely 'intelligent' action of the kind here under consideration, even among humans with the advantage of their pre-

⁹ It seems probable that when the animal meets the barrier (Sequence I, Fig. 1), R_P will be extinguished but r_X may largely escape extinction, and persevere (see pp. 238 ff.) to the point where the choice is made between the two other paths. In that event s_X would appear in the first stimulus complex (and perhaps later ones) of Fig. 3. This would give r_U , as well as r_V and r_W (Sequence II, Fig. 1), an additional excitatory tendency at the outset. In the interest of simplicity of presentation this is left out of Fig. 3, since the deduction does not require it.

sumably additional verbal-symbolic powers, constitutes substantial general corroboration of this deduction.

(2) The excitatory tendencies behind the sequence $R_R \rightarrow R_S \rightarrow R_T \rightarrow R_U$ (Fig. 4) lack the strong support of the S_D , which it has in III of Fig. 2. From this consideration follows the corollary that this particular segment of the completed act of insight should be made more slowly and haltingly, that it should be more easily over-ridden by contrary tendencies and more subject to disruption by external inhibitions, etc., than would the final section, $R_U \rightarrow R_V \rightarrow R_W \rightarrow R_X$, which has its accustomed support of S_D (Sequence II, Fig. 2).

(3) As a further corollary it may be pointed out that the particular weakness just mentioned should be especially marked at the outset of the sequence; *i.e.*, at act R_R , since this act lacks its accustomed stimulus S_D , and also the proprioceptive stimulus from a preceding act in the series such as is found elsewhere throughout all sequences which are really integrated. This, however, is also true of the first act of all series (Fig. 2).

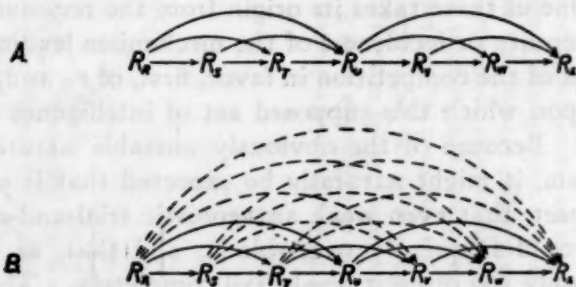


FIG. 4. Diagrammatic representation of both immediate and remote excitatory tendencies when (A) Sequences III and II of Fig. 1 first function as a (frail) unity, and when (B) they have functioned together with reward frequently enough to produce a thoroughly integrated whole with a full set of remote excitatory tendencies. The broken lines in the lower diagram represent the additional remote excitatory tendencies which constitute the basis of the more complete integration of the two habit segments supposed to come with contiguous association.

As a matter of observed fact, long series, at least, do commonly show at their beginning exactly the sluggishness which a mechanism of this kind would produce (8).

(4) As a fourth corollary, it should be pointed out that on

the average there ought to be a slight tendency to a relatively lower speed of locomotion especially in the neighborhood of R_U at the first execution of a complete intelligent act. The reason for this expectation is that since each of the two segments has been practiced as a unit, there should be remote excitatory tendencies extending throughout each series which should support the action of each segment. But, since the two segments have not been practiced in immediate succession, we should not expect such facilitating tendencies to extend across the junction from one to the other. The situation is shown in section *A* of Fig. 4. The relative lack of remote excitatory tendencies in the neighborhood of R_U should diminish the vigor of reaction at this point below what it would be in a well-integrated sequence, which, among other things, should diminish the speed of locomotion. It is to be expected, however, that this tendency may be difficult to detect unless comparison is made with a good control because the sequence beginning with R_U (Corollary 2) presumably will have a faster rate than the one immediately preceding it.

(5) After a small amount of rewarded practice, however, the act would cease to be specifically intelligent in the sense of depending upon the r_U -mechanism elaborated above, and the remote bonds would be formed as in *B*, of Fig. 4. At this latter stage there would, obviously, be no tendency to retardation in the neighborhood of R_U . Moreover, the fractional goal reaction of R_X (i.e., r_X) would become associated with the reactions of the sequence $R_R \rightarrow R_U$ as well as $R_U \rightarrow R_X$, so that what were two segments before would become in a functional sense a single segment, though doubtless in certain situations (8, 406 ff.) the components would retain a capacity to unite separately in new combinations.

(6) It is clear that the combination of two habit segments is a special case because the posterior segment has its posterior end securely anchored at the goal reaction (R_X), and the anterior segment has its anterior end securely anchored to the external stimulus of the problem situation (S_R). The more general case is that in which the solution sequence is made up from three or more segments; in the case of a spontaneous

assembly of such habit segments, the middle segment would be attached directly neither to the goal reaction nor to the external stimulus component of the problem situation. A careful examination of the presumptive behavior of the type of anticipatory mechanism such as r_U and r_R (Fig. 2) leads to the belief that such an event might possibly occur. The linkage in this case would, however, appear to be far more tenuous even than that represented in Fig. 3 and emphasized in Corollary 1 above. It is accordingly believed that such a type of spontaneous habit synthesis will be very rare indeed, even in organisms especially gifted in the functioning of anticipatory goal and sub-goal reactions.

(7) It is evident, however, that chains of three or more segments may easily be formed on separate occasions by the action of the intelligence mechanism elaborated in Section IV above, supplemented by the consolidating or integrating principle of association by contiguity which would begin operating immediately upon the initial act of insight. Thus segments 'one' and 'two' might be joined on one occasion by insight; later, segments 'two' and 'three' might be similarly joined. Now, the two newly compounded segments overlap with respect to segment 'two'; this overlapping reduces the three original segments to the functional status of two, thus permitting the insight mechanism to operate substantially as elaborated in Section IV.

VI. THE PROBLEM OF THE INITIAL SPONTANEOUS USE OF IMPLEMENTS

Special interest attaches to the results of the preceding analysis because it rather looks as if the associative mechanisms there elaborated are also responsible for the spontaneous utilization of objects as implements or tools. An instance often cited as an example of such action is that described by Köhler, on the part of his chimpanzee, Sultan (10, 132 ff.). In this case the ape had learned to drag in through the bars of his cage a bit of fruit placed a certain distance outside. On a particular trial, however, Köhler put in the cage two bamboo sticks, neither of which was long enough to reach the food.

The animal first tried to obtain the fruit with one of the short sticks. After repeated failure he began playing with the two sticks. This (apparently) random manipulation finally resulted in the insertion of the end of the smaller stick into the hollow end of the larger. Soon after this took place, the animal reached through the bars with the combination stick and obtained the fruit. This union of action segments constituted the solution of the problem. Unfortunately it is not entirely clear whether the animal had ever before joined two objects in a somewhat similar manner, and there is no control in the plan of the experiment to show that the solution was not a mere accident. To be a clear case of insight in the sense of the term here employed, the joining of the sticks should have been performed a number of times in the recent past in random play, but should never have been used in the securing of food.

An experimental arrangement which promises less ambiguous evidence of insight in the field of implement or tool using might be set up on the analogy of the Maier experiment, somewhat as follows:

Sequence I. An ape is trained to press upward, by means of a stick three feet long, against a toggle electric switch placed on the wall out of his reach; after which an electrically operated automatic machine close at hand will give him a grape.

Sequence II. A similar switch is placed three feet higher on the wall of another room, and the ape is trained to operate this with a six-foot stick which must be chosen from among two three-foot ones also available. After the pushing of the switch, a machine similar to that of Sequence I will give the animal a grape.

Sequence III. The animal is trained to put two sticks, each three feet long, together in such a way that one fits into a socket in the end of the other, making a stick six feet long. This training should be in still a different room. The reward in this case must be different from that of Sequences I and II, possibly the release from the room to return to the animal's living quarters.

Sequence IV. In a fourth room, the ape learns to place two sticks (other than those used in Sequence III) together so as to make a T, the reward in this case to be the same as that of Sequence III. One of the above sticks must have a socket in its middle instead of at its end, in order to permit the construction of the T.

The Problem. This would be set by placing the animal in the room with the high switch, as usual, but with two short sticks instead of the long one. This situation corresponds to the barrier placed in Path $R \rightarrow X$, in the rat experiment discussed above. The stick to be inserted into a socket would be like that in Sequences III and IV, but the other stick would have sockets both at one end and in the middle. *Insight would be shown by the animal tending to put the sticks together end to end rather than in the form of the T.*

A great variety of problems bearing on the nature of the presumably numerous mechanisms mediating intelligent behavior may be set up on this general pattern. In addition to the use of children at various ages, rich returns probably await the systematic utilization of such experimental approaches with the various intellectual levels of young adults in homes for the feeble-minded. It is conceivable that, in addition to filling out the picture of intelligence as delineated by investigations concerned mainly with intelligence testing, a matter of pure science and of the greatest importance, the leads so obtained might yield valuable hints for increasing the validity of the tests in actual practice.

VII. A SECOND CASE OF SUPPOSED 'REASONING' IN RATS, BY WAY OF CONTRAST

It may be well at this point to consider a second case of 'reasoning' in rats which Maier has bracketed with that discussed above on the grounds that in both, the animals are able to solve a problem by 'combining two experiences' (19). A diagram of the 'Y' maze employed in this second experiment is shown in Fig. 5. After being allowed to explore the maze for some hours, the rats were given from three to six days of 'preliminary tests,' the results of which are not reported. These may be regarded as practice or training in the act later to be tested. In this practice the animal was placed on one of the table tops, at X for example, where he found some food which he ate in part. The animal was then placed at Y and permitted to take his choice between going to Z or X. In case of the latter choice at point Q, the remainder of the food would be found and eaten. This would

involve turning to the right. At other times, with the food tasted at X as before, the animal was placed at Z and allowed to take his choice of going to X or Y at point Q, which required a *left* turn to reach the food. In a similar manner all the combinations in both directions, a total of six, would be practiced.

Then the routine or recorded tests were begun. So far as one can determine from the published report, these were

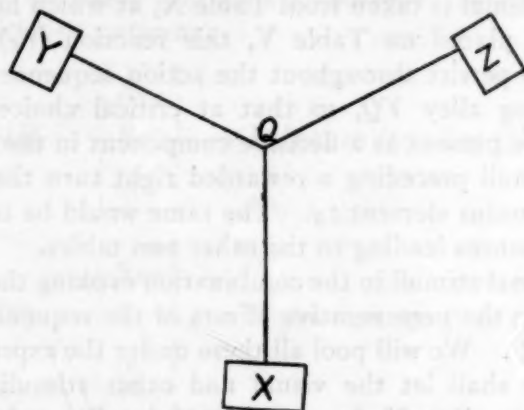


FIG. 5. Diagram of a rat maze used by Maier. X, Y, and Z are small table tops. These are connected by pathways which intersect at Q. In the present theoretical analysis, the alleys are assumed to be closed, though Maier actually used open pathways.

substantially the same as the training tests just considered. The problem was to determine whether, after such preliminary training, the animals would choose at point Q the correct one of the two possible paths more frequently than chance; *i.e.*, more than 50 per cent of the trials. Under these conditions, Maier reports in one typical experiment involving 18 trials on each of 19 normal animals (19), that 80.7 per cent of the choices were correct, *i.e.*, were such as to lead directly back to the food just tasted. How is this excess of some 30 per cent above chance to be accounted for?

It is probable that a number of different mechanisms are jointly operative in this situation. It is believed, however, that the following one, if not dominant, is a major component

and quite sufficient alone to account for all the phenomena so far reported.¹⁰ At bottom the explanatory hypothesis here to be presented is that the animal, having eaten food on a given table, *e.g.*, Table X (Fig. 5), has his body thrown into a characteristic posture which is to a certain extent distinct from that resulting from eating on any other table, and that this distinct action tendency (r_x) may persist to progressively diminishing degrees for a certain length of time. When, now, the animal is taken from Table X, at which he has just eaten, and placed on Table Y, this reaction (r_x) may be assumed to persist throughout the action sequence involved in traversing alley YQ , so that at critical choice point Q there will be present as a decisive component in the combination of stimuli preceding a rewarded right turn the proprioceptive stimulus element s_x . The same would be true of the action sequences leading to the other two tables.

Additional stimuli in the combination evoking the reaction at Q will be the perseverative effects of the sequence of acts, from Y to Q . We will pool all these under the expression S_r . Lastly, we shall let the visual and other stimuli of Q , as viewed from alley Y , be represented by S_{rq} . As a result we have at the critical choice point Q the following stimulus combination always preceding a rewarded right turn:

$$(I) \quad \begin{bmatrix} S_{rq} \\ S_r \\ s_x \end{bmatrix} \rightarrow \text{Right turn (to alley } QX)$$

¹⁰ The principles put forward in Section IV above lead to the expectation that animals would tend to go to a table on which they had just tasted food even if they had received no specific training with the tasting technique but had merely been trained to go from one table to another with a food reward. The detailed deduction is too complicated to be given here. Its substance is that this training would bring forward in the behavior sequence the fractional goal reaction, *e.g.*, r_x ; thus s_x would be present in the stimulus complex at Q much as in formula I (p. 28). When, later, the animal is given the tasting test for the first time, the perseverating s_x so aroused would summate with the s_x produced by the anticipatory mechanism. But increase in the strength of a conditioned stimulus increases the strength of the reaction tendency (23, 384). The s_x so enhanced would therefore be prepotent in evoking a choice of alley QX over any other action tendency due to the stimulus arising from an anticipatory reaction alone.

In a similar manner we have preceding the rewarded choice from *Y* to *Z*:

$$(II) \quad \begin{array}{|c|} \hline S_{YQ} \\ \hline S_Y \\ \hline S_Z \\ \hline \end{array} \rightarrow \text{Left turn (to alley } QZ)$$

Likewise the two rewarded choices, when the animal starts at *X*, will be preceded and (later) controlled by the stimulus combinations or patterns:

$$(III) \quad \begin{array}{|c|} \hline S_{XQ} \\ \hline S_X \\ \hline S_Y \\ \hline \end{array} \rightarrow \text{Left turn (to alley } QY)$$

$$(IV) \quad \begin{array}{|c|} \hline S_{XQ} \\ \hline S_X \\ \hline S_Z \\ \hline \end{array} \rightarrow \text{Right turn (to alley } QZ)$$

Lastly, the two rewarded choices at *Q* when the animal starts at *Z* will be preceded and (later) controlled by the stimulus combinations or patterns:

$$(V) \quad \begin{array}{|c|} \hline S_{ZQ} \\ \hline S_Z \\ \hline S_X \\ \hline \end{array} \rightarrow \text{Left turn (to alley } QX)$$

$$(VI) \quad \begin{array}{|c|} \hline S_{ZQ} \\ \hline S_Z \\ \hline S_Y \\ \hline \end{array} \rightarrow \text{Right turn (to alley } QY)$$

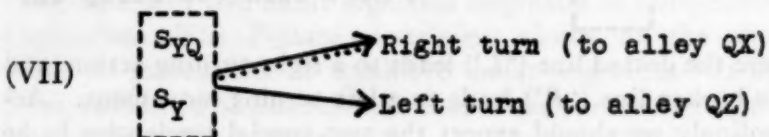
By a study of these stimulus formulæ it will be seen that there is a different pattern or combination of stimuli in each of the six cases. In particular, it is to be noted that the otherwise puzzling question of why, for example, the organism

should turn *right* at Q when going to X from Y and to the *left* when going to the same table from Z is answered by the distinctiveness of stimulus combinations shown in formulæ I and V respectively. Thus is explained the excess of 30 per cent of correct choices beyond chance as reported by Maier and cited above.

Several corollaries flow from the above deduction. In the first place, the situation in formula I, say, because two of the components (S_{RQ} and S_R) are also found in stimulus formula II which is conditioned to a left turn, still presents a picture of competition of two excitatory tendencies. It follows from this that accidental factors causing variation in the relative strength of these excitatory tendencies would frequently produce choices of the 'incorrect' path in spite of the generally dominant tendency. Secondly, since perseverations are usually short-lived and progressively diminish in strength through the passage of time, we would expect a diminution in the per cent of 'correct' choices with an increase in the time between the preliminary feeding and the test. Further, this diminution should increase *progressively* with the increase in the intervening time interval. It is also to be expected that during this period any 'extra' stimulation (23, 46) evoking more or less violent reaction in the organism would tend to break up this perseverative activity. This would weaken or abolish the special increment to the strength of s_X resulting from the perseveration of the effects of eating at X , which, in turn, would diminish the strength of the tendency to make correct choices. Moreover, this weakening, other things equal, should be progressively greater the more intense the disturbing reaction.

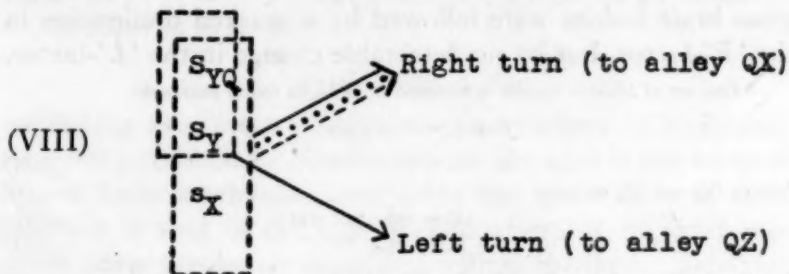
A final corollary concerns Maier's experimental distinction between the kind of behavior which we have attributed above to the operation of a perseverating fractional goal reaction, on the one hand, and what he calls ' L ' or learning, on the other. Maier produced this latter type of action by giving on a day preceding the test eight consecutive rewarded runs (with preliminary feeding, as usual) from Y to X , say. This increased amount of practice, particularly by reason of its

relative recency, would be expected to strengthen the tendency of stimulus formula I (pp. 238 f.) to evoke a right turn at Q . Assuming that this compound stimulus has not wholly crystallized into a pattern and that the individual elements possess a considerable degree of individual excitatory potentiality, it follows that S_{YQ} and S_Y , even when separated from the customary s_X , would have a considerably enhanced tendency to evoke a right turn as the result of this special training. Then, on the day of the test, even if the trial should be made without any preliminary feeding on either tables X or Z , we should have the situation represented by the following formula:



The extra shaft (dotted line) on the upper arrow indicates the enhanced strength of the tendency to turn right due to the recent practice; Maier calls this ' L ' or learning.

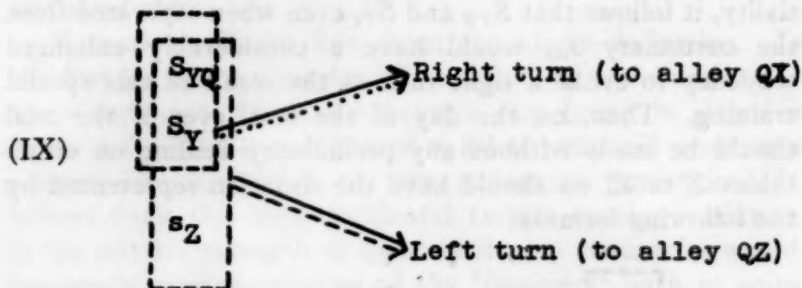
If, however, the animal has been fed at X just before the trial, this adds to the stimulus complex the element s_X which contributes another considerable element favorable to the right turn. The stimulus formula for this situation will be:



The third shaft (broken line) added to the upper arrow represents the increment to the strength of the right-turning tendency produced by adding s_X to the stimulus complex. This is the element which Maier calls ' R .' Accordingly,

under these conditions we should expect an increase in the per cent of correct choices over conditions shown in formula VII. Formula VIII represents what Maier calls ' $R + L$.'

But if, instead, the animal is first allowed to taste on table Z, the stimulus formula will be:



Here the dotted line (' L ') leads to a right-turning action, and the broken line (' R ') leads to a left-turning movement. Accordingly we should expect the two special tendencies to be opposed, *i.e.*, the situation is what Maier has called ' $R - L$.' Clearly, then, stimulus formula IX should give a smaller percentage of reactions of the kind associated with the perseverating reaction than should stimulus formula VIII.

Once experimental values are available for ' $R + L$ ' and ' $R - L$,' there may be computed from them approximations to the values of ' R ' and ' L ' separately.¹¹ By means of this combination of techniques, Maier reports that, on the average, gross brain lesions were followed by a marked diminution in the ' R '-factor, but by no detectable change in the ' L '-factor.

¹¹ One set of Maier's results is reported to yield, in round numbers:

$$R + L = 89 \text{ per cent} \quad (1)$$

$$R - L = 70 \text{ " " } \quad (2)$$

Adding (1) and (2), we have

$$\begin{aligned} 2R &= 159 \text{ per cent} \\ R &= 79.5 \text{ " " } \end{aligned} \quad (3)$$

Substituting (3) in (1), and solving for L , we have

$$L = 9.5 \text{ per cent}$$

Under these conditions ' R ,' or the perseverational factor, turns out to be about eight times as potent as this amount of special training on ' L .' It seems likely that this general approach to such complex problems is capable of much wider use than is the case at present.

In this way, assuming that the experiment substantially survives repetition, Maier appears to have proved effectively both the reality and the distinctiveness of the two mechanisms or principles of action quite as is to be expected on the basis of the preceding analysis. But to have shown that the two processes, 'R' and 'L', are genuinely distinct is by no means to have shown that 'R' in the Y-maze situation just considered is the same in its essential structure as what Maier calls reasoning in the situation described in Section II above, much less that it is a case of reasoning in any generally recognized sense. The results from the experiment discussed in Sections II to IV have the appearance of a spontaneous combination of two habit segments originally of independent acquisition. The Y-maze experiment shows, on the other hand, no indications of combined habit segments. It appears, rather, to be essentially a case of the perseveration of a goal reaction as a pure-stimulus act from one stimulus situation to a closely following one, the proprioceptive stimulus from this perseverating goal reaction becoming the critical element in differential stimulus compounds or patterns. Indeed, if the hypothesis of a perseverative goal reaction should prove to be sound, it seems doubtful whether it would be legitimate in this second case to designate (1) the preliminary eating, and (2) the conditions of the subsequent choice at Q, even as separate 'experiences'; such a continuity would seem to make them one.

VIII. SUMMARY

Taking as a point of departure an analysis of intelligent behavior published by Norman Maier, the view is put forward that a basic mechanism mediating one great class of such behavior is that of the spontaneous assembly of habit segments never previously associated with each other. Behavior of this kind is deduced from stimulus-response principles. If this deduction should turn out to be sound, it would amount to a proof that at least one form of insightful behavior is not an ultimate, unanalyzable entity but is, instead, a special, though somewhat complicated, case of association between

stimuli and reactions. Indeed, the principles of action which appear by this deduction are of such a nature that they might conceivably be incorporated into an inorganic machine which would automatically solve such problems. This is not to deny either the reality or the importance of such forms of behavior. On the contrary, it leads to the view, mainly on the grounds of the novel character and distinctiveness of the mechanism mediating it, that an adequate understanding of such behavior is of special importance, both theoretical and practical. The above analysis does lead, however, to the denial of the necessity of postulating any peculiarly experiential, psychic, or configurational factors in order to explain the existence of this particular form of intelligent action.

Whether the forms of behavior involving the spontaneous assembly of habit segments shall be called 'reason' as Maier and Shepard are inclined to do, is not primarily a question of science, but, rather, one of taste and convenience. The present writer inclines to reserve the term 'reason' for application to those problem situations in which the solution is mediated through the action of symbolic verbal reactions. There seems even less justification for applying the term 'reasoning' to the behavior of animals on the Y-maze under the tasting technique which we have explained on the basis of perseveration.

REFERENCES

1. BAERNSTEIN, H. D., AND HULL, C. L., A mechanical model of the conditioned reflex, *J. General Psychol.*, 1931, 5, 99-106.
2. BENNETT, G. K., AND WARD, L. B., A model of the synthesis of conditioned reflexes, *Amer. J. Psychol.*, 1933, 45, 339-342.
3. HOTELLING, H., Analysis of a complex of statistical variables into principle components, *J. Educ. Psychol.*, 1933, 24, 417-441; 498-520.
4. HULL, C. L., Knowledge and purpose as habit mechanisms, *PSYCHOL. REV.*, 1930, 37, 511-525.
5. —, Goal attraction and directing ideas conceived as habit phenomena, *PSYCHOL. REV.*, 1931, 38, 487-506.
6. —, The goal gradient hypothesis and maze learning, *PSYCHOL. REV.*, 1932, 39, 25-43.
7. —, The concept of the habit-family hierarchy and maze learning, *PSYCHOL. REV.*, 1934, 41, Part I (Jan.), 33-52; Part II (March), 134-152.
8. —, The rat's speed-of-locomotion gradient in the approach to food, *J. Comp. Psychol.*, 1934, 17, 393-422.

9. KELLEY, T. L., Crossroads in the mind of man, Stanford Univ. Press, 1928.
10. KÖHLER, W., The mentality of apes, Harcourt Brace and Co., 1925.
11. KRIM, N. B., Electrical circuits illustrating mammalian behavior and their possible engineering value, Thesis presented for the degree of Bachelor of Science, Mass. Institute of Technology, 1934; copy deposited in the Yale Library.
12. KRUEGER, R. G., AND HULL, C. L., An electro-chemical parallel to the conditioned reflex, *J. General Psychol.*, 1931, 5, 262-269.
13. LIPPMANN, W., A future for the tests, *The New Republic*, 1922, 33, 9-11.
14. LUMLEY, F. H., An investigation of the responses made in learning a multiple choice maze, *Psychol. Monog.*, 1931, 42, No. 189.
15. —, Anticipation of correct responses as a source of error in the learning of serial responses, *J. Exper. Psychol.*, 1932, 15, 195-205.
16. —, Anticipation as a factor in serial and maze learning, *J. Exper. Psychol.*, 1932, 15, 331-342.
17. MAIER, N. R. F., Reasoning in white rats, *Comp. Psychol. Monog.*, 1929, July, No. 29.
18. —, Reasoning and learning, *Psychol. Rev.*, 1931, 38, 332-346.
19. —, The effect of cortical destruction on reasoning and learning in white rats, *J. Comp. Neurol.*, 1932, 54, 45-75.
20. MILL, J. S., A system of logic (8th edition), Harpers, 1887.
21. MILLER, N. E., A reply to 'Sign-Gestalt or Conditioned Reflex?' *Psychol. Rev.*, 1935, 42, 280-292.
22. MITCHELL, M. B., Anticipatory place-skipping tendencies in the memorization of numbers, *Amer. J. Psychol.*, 1934, 46, 80-91.
23. PAVLOV, I. P., Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex (trans. and ed. by G. V. Anrep), London: Oxford Univ. Press, 1927.
24. SPEARMAN, C., Abilities of man, N. Y.: The Macmillan Co., 1927.
25. SPRAGG, S. D. S., Anticipation as a factor in maze errors, *J. Comp. Psychol.*, 1933, 15, 313-329.
26. —, Anticipatory responses in the maze, *J. Comp. Psychol.*, 1934, 18, 51-73.
27. THORNDIKE, E. L., Animal intelligence, N. Y.: The Macmillan Co., 1911.
28. THURSTONE, L. L., The theory of multiple factors, 1932, Edwards Bros., Ann Arbor, Mich.
29. WOLFE, J. B., AND SPRAGG, S. D. S., Some experimental tests of 'reasoning' in white rats, *J. Comp. Psychol.*, 1934, 18, 455-469.

[MS. received November 9, 1934]

SCIENTIFIC MEASUREMENT AND PSYCHOLOGY

BY DOUGLAS MCGREGOR

Harvard University

Measurement is basic to every scientific discipline. Philosophers, aware that the value of any scientific fact is dependent upon the validity of the method of measurement by which it has been established, have always been concerned with scientific measurement. Of recent years scientists themselves have become cognizant of the problems involved. As a result, many treatises dealing with various aspects of the subject have appeared. The new approach to scientific problems associated with Einstein's name has supplied impetus to the renaissance of scientific philosophy.

All this recent concern about the validity of scientific measurement has aroused but slight interest among psychologists, possibly because they think that a great gulf exists between physical and psychological methodologies. It is the purpose of this paper to demonstrate that such a gulf does not exist, and to present an interpretation of the systematic importance of modern theories of measurement for psychology.¹ If we, as psychologists, are to have measurement of any value whatever we must accept the logical restrictions imposed within other sciences or devise a separate valid system for ourselves. We claim to use measurement in psychology, for we perform experiments which depend on measurement and we express our results in the form of tables, graphs, laws, and

¹ Mr. McGregor has very kindly offered me this footnote in order that I may say that he and I have worked on this paper in close collaboration (although he has done the hard work and I have only criticized) and that I am in hearty accord with everything that he says. I am convinced that the difficulties of psychology in formulating the problems of mental measurement have been mostly the result of too long an adherence to the now outmoded Cartesian dichotomy of mind and matter. Thus I think that the author has not only made it clear that psychologists frequently measure their phenomena by valid scientific processes, but that he has also contributed importantly to systematic psychology in an argument which I fully endorse and hope may prevail.—E. G. BORING.

statements about functional relationships. We are accustomed to discuss the theoretical implications of these results with perfect confidence, although we usually ignore the obvious fact that unless the methods of measurement are valid, the conclusions drawn may be meaningless. In a consideration of this scientific situation we must begin with an examination of the logical requirements for a valid methodology of measurement.

THE BASIS OF MEASUREMENT

The concept of relativity and the name of Einstein have gone hand in hand for two decades, and the interpretation of scientific fact which gave rise to the famous transformation equations has vitally affected scientific thinking. Einstein demanded that the meaning of the concepts of length and simultaneity, which purport to apply to concrete physical situations, should be sought in the definite physical operations involved in their application. Physicists² and logicians³ have been responsible for the general application of this point of view to scientific methodology. Operationism⁴—the belief that an entity is adequately defined only in terms of the specific operations involved in its observation—is an important premise in most recent works devoted to measurement.

Let us consider length as a specific example illustrating the operational procedure. Its meaning is to be sought in those operations by which the length of a concrete physical object is determined. The operations for the determination of the length of a physical object vary with the order of magnitude of the object. Microscopic and macroscopic length are two different magnitudes, although closely related because operations can be found which are common to both concepts. With the ordinary scale the operation is the judg-

² P. W. Bridgman, *The logic of modern physics*, New York, 1928. N. R. Campbell, *Physics, the elements*, Cambridge, England, 1920.

³ M. R. Cohen, and E. Nagel, *An introduction to logic and scientific method*, New York, 1934.

⁴ All recent writers on measurement make fundamental use of the technique that Bridgman has called *operational*, so that we can take this neo-positivistic view as something new, an 'operationism.'

ment of the coincidence of the ends of the object with marks on the scale; a matter of observer-discrimination in respect of the congruence of two visual extents approximately superposed.⁶

The physicist is well aware that the basis of all measurement is the psychological judgment, the operation of discrimination. Even the simplest form of measurement, enumeration, involves this operation. However, the physicist tends to limit the kinds of discrimination involved to those for which universal agreement can most closely be approximated. The fundamental discriminations are these:⁶

1. Simultaneity, consecutiveness, and 'betweenness' in *time*.
2. Simultaneity, separateness, and 'betweenness' in *space*.
3. Number.

Ultimately these, and all other discriminations involved in measurement, are reduced in effect to the discrimination of difference. The judgment of equality does not immediately fit this classification, but analysis of the psychological judgment of equality from an operational point of view (the analysis of the actual process by which one arrives at a particular judgment *equal*) readily shows that equality can only be defined negatively in terms of the inability to discriminate. For example, an object is said to be 10 centimeters long when, under specified conditions, the observer can detect no difference in length between the object and a portion of the scale comprising 10 centimeters.

The instability of the 'equality' judgment is well known in psychophysics, so that even in psychology the phenomenal experience of equality has become suspect, and equality is being defined—operationally—as a point of symmetry between two differences, 'greater' and 'less.'⁷ The operation is one of equalization.

⁶ Measurement of length is perfectly feasible when the object and the scale are separated, but the variability of the operational judgment is increased. The operation actually used is chosen because it provides the greatest consistency, the least 'error' of judgment.

⁷ N. R. Campbell, *op. cit.*, p. 29.

⁸ S. W. Fernberger, *PSYCHOL. REV.*, 1930, 37, 107-112.

The operational technique, then, and, in particular, the discrimination of difference, are basic to modern treatments of measurement. The importance of these ideas will become increasingly evident in the following analysis of the different forms of measurement.

ENUMERATION AND ORDER

Science deals primarily with quantities which can be expressed by means of numbers. Measurement can be defined as 'the process of assigning numbers to represent quantities.'⁸ Numbers first enter as a means for replacing with exact statements the less satisfactory general impressions of quantity. Enumeration is undertaken because significant relations are suspected between the groups counted.

If enumeration is to be satisfactory, we must first have sufficient knowledge about a subject to distinguish it from other subjects. If we wish to count the number of drops in a glass of water, we must define a glass of water, and a drop. These two definitions involve discrimination. A further logical requirement is that the population of a group must be an invariant property of the group; that is to say, the population must be independent of the person counting and of the order in which the group is counted.

An important requirement for satisfactory enumeration arises in the logic of mathematics. It is the requirement of order. Order is based on the logical relations of *transitivity* and *asymmetry*. The relation of transitivity can be expressed symbolically thus: If $a-R-b$, and $b-R-c$; then $a-R-c$, where the symbol R is read: 'has the relation R to.'⁹ Asymmetry may be expressed thus: If $a-R-b$; then $b-\text{not } R-a$.¹⁰

The operations for the demonstration of these relations are simple. Let us take two groups of discrete objects. We remove one object from each group alternately until we have removed all the objects from one group. If the last object removed follows directly upon the removal of the last object

⁸ Cohen and Nagel, *op. cit.*, p. 289.

⁹ E.g., if a is greater than b , and b is greater than c , then a is greater than c .

¹⁰ E.g., if a is greater than b , then b is not greater than a .

in the group with which the process was started, then the groups are defined as equal, and the relation of symmetry is fulfilled. On the other hand, if we are left with a remainder in one group after completion of the cycle containing the last member of the other group, the group having the remainder is defined as greater than the other, and we have demonstrated the relation of asymmetry. With three groups and a similar process we can demonstrate the relation of transitivity.

INTENSIVE MAGNITUDES

However, measurement is seldom a mere matter of simple enumeration. Many characteristics of objects cannot be sharply distinguished, and hence cannot be counted, because they seem to form a continuous series. One cannot discriminate adequately the sharpness of knives, the hardness of minerals, or the alertness of children without further definition and qualification of these properties. When we wish to know the degree of such properties, we undertake to assign numbers to indicate different degrees, so that we may count with reasonable precision. Such measurement produces three distinct kinds of magnitudes depending upon the logical requirements which can be operationally satisfied.

The simplest of these are called *intensive magnitudes*. For them the relations of transitivity and asymmetry can be established, but little else. A common example of an intensive magnitude is hardness of minerals. The discriminative operation in this case is couched in terms of the ability of one mineral to scratch others. One mineral is defined as harder than another if it will scratch the other, and an observer can say which scratches which. Thus we can arrange minerals in an order on the basis of their ability to scratch one another, and then we can assign numerals to the members of the series. In this manner hardness is said to be measured intensively.

It is essential that the relations of transitivity and asymmetry be fulfilled, for if the relation were symmetrical we should find that a mineral *A* would scratch a mineral *B*, and that *B* would also scratch *A*. Then there would be no reason for putting one before the other in the series. If the relation

were not transitive *A* would scratch *B*, and *B* would scratch *C*; yet *A* would not scratch *C*, but *C* would scratch *A*. We could not say then whether *C* belonged before *A* or after *B*.

A body can be found which neither scratches nor is scratched by some other body in the series. It does not fulfill the conditions for a series, but it may behave toward all the other members of the series exactly as does the body which it does not scratch. In this case the two bodies which do not scratch each other are assigned the same numeral and defined as equal in hardness.¹¹

Hardness, as a magnitude, is distinctly limited in its usefulness. There is no way to demonstrate that the ratio of the numerals assigned to different bodies is the same as the ratio of the degrees of hardness. Although to diamond is assigned the numeral 10, to ruby 9, and to talc 1, the difference between diamond and ruby is nowadays believed to be much greater than the difference between ruby and talc. The order of the series is fixed by experiment, but, when we have numbered one body 8, there is no means for showing whether the next harder member of the series should be 9, or 30, or 100. With weight, on the other hand, when we represent one body by the numeral 8 and another by 3, we can demonstrate that the ratio of the two weights is exactly the same as the ratio of the numbers.

With all intensive magnitudes we are limited to statements of *greater* or *less*. We cannot validly undertake to state *how much greater* or *how much less*.

It will be noted that the psychologist has to use many intensive magnitudes. For instance, an observer can arrange various substances in an ordered series on the basis of their saltiness, and we can establish saltiness as an intensive magnitude like hardness. The relations of transitivity and asymmetry are fulfilled for the series, but we cannot state that one substance is three times as salty as another until we find a further operation that gives meaning to this statement.

Other examples of intensive magnitudes are pleasantness

¹¹ Equality is here defined in the negative fashion previously suggested. The experimenter defines two bodies as equal in hardness only when he is unable to observe that either body has scratched the other.

(as derived by the method of paired comparisons), sensory intensity (when it is defined in terms of the summed number of j.n.d.'s above the threshold), and personality traits (as determined by the various tests).

FUNDAMENTAL MAGNITUDES

With hardness, letters of the alphabet (since their order is fixed) would serve just as well as numerals to distinguish members of the series. When a symbol has been chosen for one mineral we are not limited, *except as to order*, in our choice of a symbol for any other mineral. With weight, on the other hand, when the number for one weight has been selected, there is no choice left for the other members of the series. The reason for this important distinction is that weight not only satisfies the relations of transitivity and asymmetry, but is also *additive*. Magnitudes for which additivity can be demonstrated are called *fundamental*.

Additivity is basically a mathematical problem. The development of arithmetic is dependent upon a few logical propositions which determine the nature of addition.¹² The physicist accepts these propositions, taking the position that any property which can be shown operationally to obey them is additive.

The 'First Law' of addition¹³ states that $a + 1 > a$, where 1 may be any amount of the property greater than the experimental error. The operation of addition, with weight, consists in placing two bodies together on one scale pan of the balance. We shall also require the operation of equalization, which may be defined as follows: Two bodies are of equal weight if, when they are placed successively in the same scale-pan, neither of them disturbs the balance with a constant weight in the other pan.

It is obvious that we could assign no numeral but 1 if the First Law were untrue. As a matter of fact, when a property does not obey the law, every effort is made to define it so that

¹² A. N. Whitehead and Bertrand Russell, *Principia mathematica*, Cambridge, England, 2d ed., 1927.

¹³ Campbell, *op. cit.*, p. 280.

$1 + 1 = 1$. For example, though one may add quantities of water, the density remains 1.

The next step in establishing a fundamental magnitude is to devise a standard series in terms of which other bodies can be measured. Let us illustrate this process in the case of weight.

First, we select arbitrarily any body, B_1 , for the unit. Then, using the balance and the operation of equalization, we find $B_1' = B_1$. Now, using the balance and the operations of addition and equalization, we add B_1' and B_1 (by placing them in one scale-pan), and find $B_2 = B_1' + B_1$. Then we place B_2 and B_1 in one pan, and find B_3 equal to their sum. Then we find $B_4 = B_3 + B_1$, and so on. Continuation of this process provides a standard series fulfilling the requirements of the laws of addition for weight. We shall actually need but a few members in the standard series. For example, the bodies 1, 2, 2, 5, 10 will measure any integral body up to 20.

The 'Second Law' of addition states: "The magnitude of a system produced by the addition of bodies $A, B, C, \dots n$ depends only on the magnitude of the bodies, and not on the order or method of their addition."¹⁴

This Second Law is nearly always true if the First Law is true, but there are important exceptions. If both laws of addition cannot be demonstrated, the magnitude cannot be validated as additive, and a third form of measurement must be relied on. Quantity of heat is an example of a magnitude for which the Second Law fails. The quantity of heat in a body A is defined as greater than that in B , if A , dropped into a certain volume of water, raises its temperature more than B , dropped into the same volume. The First Law is fulfilled, for a hotter body dropped into a colder liquid always raises the temperature of the liquid. However, the Second Law is un-

¹⁴ Campbell, *op. cit.*, p. 283.

There are several corollaries of this law:

equals + equals = equals,

$a + b = b + a$ (The commutative law of mathematics),

$a + (b + c) = (a + b) + c$ (The distributive law of mathematics),

if $a + b \geq c$; then $b + a \geq c$,

if $(b + c) \geq d$; then $a + (b + c) \geq a + d$.

fulfilled, for equals added to equals do not always give equals. The reason is that bodies have a finite heat capacity and do not give up all their heat to the water. Quantity of heat is, therefore, measured by a process which we shall discuss shortly.

Multiplication of additive magnitudes can be operationally established by a process of repeated addition. Fractional magnitudes can be established by a conceptual change of unit. If we choose a new unit, say one-half the size of the old, then the number representing the measure of any concrete instance of the magnitude in terms of the new unit will be twice as large as the number which was its measure in terms of the old unit. In general: If we change the size of the unit by any factor, we can change the measure of any concrete instance by the reciprocal of that factor, and so leave unchanged the ratio of the measures of any two instances.

Fractional magnitudes depend upon the validity of the Second Law. When only the First Law is established, one cannot prove that $5 + 7 = 12 = 3 + 9$, nor can one show that $1/2 = 5/10 = 25/50$.

DERIVED MAGNITUDES AND NUMERICAL LAWS

Measurement, as we have considered it, is only a means to an end. Having established fundamental magnitudes, science is concerned to discover how various properties are related to one another. The assertion of a relationship between two or more magnitudes is a *numerical law*. Some, but not all, numerical laws measure what are termed *derived magnitudes*.

When two properties appear to hold a definite relationship, the one to the other, the usual procedure is to observe a large number of correlated values of both, and to write these down in parallel columns. Then a relationship which can be expressed as a generality is sought between the columns. For example, the statement that a magnitude A varies inversely with another magnitude B is the statement of a numerical law relating A and B .

Very often such relationships can be expressed in the form $y = f(x)$. The equality sign in an equation of this sort is no

more than an indicator of a relation between the two parts of the equation. It is not a statement of a symmetrical relation. The expression $y = f(x)$, used in this way, is a statement that certain values of one magnitude are characteristically associated with certain values of another magnitude. It is also a statement that a numerical relation of the kind characteristic of mathematical variables holds between these associated values.

When we measure different volumes of a substance, and the masses associated with these volumes, we find a relation which may be expressed as $m = kV$, where k is constant. If we follow the same procedure for a different substance, we find that the relation is still true, but that k has another numerical value. We can order a series of substances in respect of their different numerical values of k .

It is also possible to measure another property of substances (termed *density*) in respect to the capacity of bodies to float or sink in liquids. We can then rank the substances in an order satisfying the logical relations of transitivity and asymmetry, i.e., we can measure density by an intensive method. We find upon inspection that the order of the substances measured by this intensive method is the same as the order of the numerical values of k , when the mass and volume of the same substances are measured. It is therefore possible to say that the constant k measures density.

Density differs from most fundamental magnitudes in that it is not additive. On the contrary, it is defined so that $1 + 1 = 1$ for a given substance. Nevertheless density is an important magnitude. Because it is difficult to measure volume by a fundamental process, we usually employ this relation between volume and mass to define volume as a given weight of a substance of a given density.

The concern of the psychologist with these functional relationships of the form $y = f(x)$ shows that the derived magnitudes are an important part of his subject-matter. For example, we measure auditory loudness as a function of frequency and intensity, intelligence as a function of mental age and chronological age, learning as a function of errors and

repetitions, and chromatic saturation as a function of intensity and composition of radiant energy. We shall consider such measurement critically in one of the following sections.

It would appear that the only real difference between fundamental and derived magnitudes is their susceptibility to addition, but even this distinction proves not to be universal because there are some derived magnitudes which can be measured by a fundamental process. As a matter of fact only a few physical magnitudes are measured in practice by a fundamental process. They are length, weight, period of time, and electrical resistance. Some of the magnitudes which are measured in practice by a derived process, but which can be measured by a fundamental process and proved additive, are the following: angle, area, volume, energy, mass, momentum, uniform velocity, intensity of radiation, electric capacity and current, potential difference, conductance, charge, and magnetic flux. The most important magnitudes which can be measured only by a derived process are uniform acceleration and temperature.

Examination of this list confirms the statement that there is no important distinction between derived and fundamental magnitudes. Magnitudes as similar as acceleration and velocity appear in different classes, while such representative derived magnitudes as resistance and uniform velocity, the very meanings of which are intimately associated with numerical laws, appear as capable of addition.

The statement that the constants in numerical laws are magnitudes is true only when the numerical laws involve but one constant. Otherwise it may or may not be true. Consider the law expressing the relation between the length of a metal rod and its temperature: $L = a + bt + ct^2$. In this law a , b , and c are all constants, but they are not all magnitudes. When we vary the constants by varying the body on which the experiments are made, we find that we cannot vary them independently; when one constant varies, so in general do the others. Accordingly, as far as we know, all the constants are associated with one property. A property which is a magnitude cannot be associated with more than one numeral, for the

magnitude must have an order, and this order cannot agree with more than one set of numerals unless all the sets of numerals have the same order. In this case a , b , and c do not have the same order. For one rod, b may be greater than c ; for another, c may be greater than b . Hence b and c cannot measure the same magnitude, and, since there is no evidence that more than one magnitude is concerned, they cannot measure magnitudes at all.

A law of this nature is termed an *empirical law*. It is not an ultimate nor a satisfactory law. Experimentation is continued in the expectation that the constants may be reduced so as to be independently variable, yielding a true numerical law.

The establishment of numerical laws usually involves graphs. If the curve of the equation is a straight line, the problem of the law is solved. If it is not rectilinear, we can proceed by trial and error. For example, if we suspect a parabolic function, we square one set of values and note whether the result is a straight line. Smooth curves are expected, partly because of the general confidence in simplicity. It is also true, in general, that, the less smooth the curve, the larger the number of constants required to express the law; and, the more constants there are, the less is the probability that all of them measure magnitudes.

In making measurements we always intend to change only two properties of a system, leaving the others unchanged, and we shall expect regularity of results only when we have succeeded. If we obtain a smooth curve, we do not hesitate to interpolate, for, if the points fixed are sufficiently numerous, it is impossible to draw through them a curve which is, at the same time, smooth and so different from any other smooth curve through these points that the difference can be detected by experiment.

DIMENSIONS

The concept of *dimensions*, leading to the intricacies of modern dimensional analysis, lies outside the realm of this paper. However, there is one kind of magnitude, important to the discussion which follows, which cannot be satisfactorily discussed without reference to the concept.

The origin of the term *dimension* is geometrical. A plane surface is said to have two dimensions; a line has one. The magnitude connected with a plane surface (area) bears to the magnitude connected with a line (length) the relation which is expressed by: $\text{area} = \text{length}^2$. The same reasoning, complicated by considerations regarding units, leads to the statement that density has the dimensions $\text{mass} \times \text{volume}^{-1}$, i.e., mass/volume . The dimensions of a derived magnitude are always expressed as a product of powers. Fundamental magnitudes are given the dimension 1 in terms of themselves.

There are derived magnitudes with zero dimensions. These occur when a numerical law states a relation between two fundamental magnitudes of the same kind, measurable in terms of the same units. For example, the law expressing the relation between the length of sides and diagonals of squares is stated: $\text{length of side/length of diagonal} = .707$. The constant .707 is a no-dimensional derived magnitude. π is likewise a no-dimensional magnitude. So are the Weber fraction ($\Delta R/R$) and the *IQ* (mental age/chronological age).

PSYCHOLOGICAL MEASUREMENT

We have now considered in broad outline the logical system of measurement constructed by modern physicists. It is evident that operationism takes the observer for granted, maintaining that selective discrimination is the operation basic to all measurement. With weight, a fundamental magnitude, the discrimination involves the pointer and scale of a balance. With length, another fundamental magnitude, the discrimination involves the congruence of superposed visual extents. With density, measured by an intensive process, the discrimination is in terms of 'floatation.' With hardness it is in terms of 'scratch.' Even with 'robot' measurement, such as that provided by the photo-electric cell, calibration is necessary, and calibration involves discrimination.

We have now to consider the use that experimental psychologists can make of measurement with the data and means at their disposal. We must continue to work with

operationally defined concepts and in terms of discrimination, for these notions are fundamental to the logic of all measurement. In such terms, however, we shall be able to understand much more exactly than ever before the nature of these 'psychological' magnitudes. If, in the process, we seem to lose the distinction between psychological and physical magnitudes, between psychology and physics, we can console ourselves by remembering that this fusion is the inevitable result of the operational technique which defines innumerable entities by reference to innumerable operations, instead of seeking first to impose the Cartesian dichotomy of mind and matter upon all reality. Strange as it may seem, this loss of psychology in science has been predestined ever since Wundt confirmed psychology as a science.

INTENSIVE MEASUREMENT

Visual brilliance provides an example of a magnitude which can be measured intensively. The operation is the discrimination of *different* in respect of two juxtaposed visual fields of the same wave-length composition. Discrimination is, of course, possible between visual fields of different wave-length composition, but, because the discrimination is poorer under these conditions, the wave-length difference is reduced in practice to zero, or nearly to zero. The operation of equalization is, as has been pointed out before, an expression of the inability of the observer to discriminate between the two fields in respect to brilliance. A second operational procedure, yielding substantially the same results, is in terms of discrimination of the presence of flicker in the visual field, under the conditions of flicker photometry. The discrimination in both cases is essentially the same as the discrimination of 'scratch' for hardness. By either photometric method we can establish an order of intensities of radiant energy satisfying the relations of transitivity and asymmetry. Brilliance is thus seen to be an intensive magnitude measurable by an operational procedure.

Chromatic saturation, loudness, weight, pressure, sweetness, pain—these and many other magnitudes can be meas-

ured by an intensive process in much the same way. The demonstration of the existence of an intensive magnitude depends only upon the capacity for selective discrimination on the part of the human or animal organism. The number of such magnitudes is limited only by the capacity of the organism to provide differentiated processes consequent upon stimulation, and its capacity for discriminating among these processes in overt response. Undoubtedly future research may bring to light new intensive magnitudes which are not recognized at present, but which the capacities of the organism render discriminatively possible.¹⁸

FUNDAMENTAL AND DERIVED MEASUREMENT

As we have already noted, many magnitudes can be measured either by a fundamental or by a derived method. Length, for example, is measured by a fundamental method. Let us consider some of the conditions under which it is measured.

Case I: It is necessary, first, to define the operations to be used in measurement. The operation of addition shall consist in placing the linear extents to be added end to end in the same plane in prolongation of each other. We shall define the operation of equalization in terms of the point of symmetry between *greater* and *less* when the two extents are parallel and close together, with one of the ends of each extent in the same perpendicular plane:

Thus defined, length is additive and fundamental. This is the example discussed in the first section of this paper.

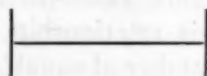
Case II: We shall define the operation of addition as in case I, but the operation of equalization shall be with the lines displaced thus:

¹⁸ S. S. Stevens, for instance, has been able to provide conclusive evidence as to the existence of tonal volume and tonal density as intensive magnitudes. See *Amer. J. Psychol.*, 1934, 46, 397-408; *J. Exper. Psychol.*, 1934, 17, 585-592; *Proc. Nat. Acad. Sci.*, 1934, 20, 457-459.

In this case we shall obtain approximately the same values as in case I, but the probable error will be larger. The reason is plain: discrimination is poorer.

If we now bring the two extents closer together we find that the probable error decreases as we approach more nearly the operational conditions of case I. If s be some measure of the displacement of the lines under the conditions for equalization, and if L_1 and L_2 be length as measured under the conditions of case I and case II respectively, then we shall find: $L_2 = f(L_1, s)$. The equation gives a measure of the relation of discriminatory capacity under the two given operations.

Case III: Let us use the same operation of addition as in previous cases, but for the operation of equalization let us add perpendicular *Nebenreize* of varying length to one of the extents, keeping their relative positions somewhat as in case II:



Here we have an illusory effect which is a function of the length of the *Nebenreize* and of the displacement of the lines. If the operation of equalization is the same as that for case I, the illusion is inoperative regardless of the *Nebenreize*. Under the conditions of case II (provided we hold the relative separation of the extents constant) we get a relation of the form $L_3 = f(L_2, d)$, where d is the length of the *Nebenreize*. The equation, which gives a measure of the illusion, is really a transformation equation showing the relation between two fundamental measurements of the same linear extents.

These last conditions of measurement (case III) provide an interesting analogy with the Fechner function. The Fechner function represents the relation between two magnitudes: $S = f(R)$. It has long been argued that S , as the sum of a series of *j.n.d.*'s, is not additive and hence is not a fundamental magnitude. The difficulties to be faced in proving *j.n.d.*'s additive are great, but their use can be avoided.

Supraliminal equal 'sense-distances' can be directly de-

terminated so as to provide a series of arbitrary 'unit sense-distances' which are defined as equal by the operation under which they are obtained. In order to demonstrate that these 'sense-distances' provide us with a fundamental magnitude, we must prove that the two laws of addition are fulfilled. Demonstration of the First Law is easy, but demonstration of the Second Law presents methodological difficulties, particularly when the 'sense-distances' are not coterminous. Nevertheless, through the use of a method of substitution similar to that employed by the physicist in establishing a standard series of weights, we can demonstrate that the sum of a series of 'sense-distances' is independent of the order of their addition. This method of substitution enables us to define the operation of addition much as it is defined for length in case I.

We can now obtain a relationship between S , defined in terms of these unit 'sense-distances', and R , defined in the usual way. This relationship can be expressed: $S = f(R)$, where S is the number of equal 'sense-distances' counted from a given origin and R is the stimulus-distance measured from the same origin. However, the more general relationship upon which this function depends is: $s = f(r, R)$, where s is any 'sense-distance,' r is any corresponding stimulus-distance, and R is the total stimulus magnitude ($= \Sigma r$). This form of expression is analogous to case III where we found $L_s = f(L_2, d)$. In fact we can regard the Fechner function as a contrast illusion. The 'sense-distance' (s) varies directly with the stimulus-distance (r), but inversely with the magnitude of the intensities (R) that limit at either end the r in question. R is the *Nebenreiz* to the distance which it limits.

The Fechner function therefore turns out to be analogous to this contrast illusion of visual length. A 'sense-distance' limited by large intensities is smaller than one limited by small intensities, just as a line limited by large *Nebenreize* is smaller than one limited by small *Nebenreize*. We have understood $S = f(R)$ and yet have avoided the Cartesian dichotomy. It is true that we have measured distances under a different set of operational conditions than those chosen by the physicist,

but our measurement is no less 'physical' because of that. What we *have* done is to shift the focus of attention from the measured *magnitude* to the operating *organism*. We have examined the relationship between a magnitude measured in the usual way and one measured under special operational conditions which are specially designed to shed light upon the functioning of the organism. But *S* is as physical, or as psychological, as *R*.

The Weber function is another example of this 'focusing of attention' upon the functional nature of the organism. We measure here the observer's discriminative capacity with respect to a given magnitude. We obtain a no-dimensional derived magnitude. The Weber function is not a true numerical law because it is found that $\Delta R/R \neq k$; but the function is an empirical law because $\Delta R = f(R)$, and probably because of something else too.

The reduction of the empirical Weber function to a true numerical law is greatly to be desired. Such an achievement would render sensitivity a function of more than one definite magnitude, and would substitute for a vague statement of capacity a knowledge of an intelligible relationship.

The photometric visibility curves provide useful examples of derived measurement. The visibility function is a plot of the reciprocal of the radiant intensity required, at different wave-lengths, to evoke equal brilliance. A visibility curve represents the relative capacities of unit intensities, at different wave-lengths, to evoke vision. Let us compare this method of derived measurement with that used for the measurement of density.

When we take different masses of a given substance and measure their respective volumes, we find that these two properties are related by a function of the form $m = f(V)$,—actually $m = kV$. This function holds for any substance, but the constant k has different values for different substances.

To measure the visibility function we choose a particular brilliance, which we define by a photometric method. Then we measure, for many different wave-lengths (λ) of spectral light, the radiant intensity (I) required to bring that color up

to the chosen brilliance. We find a functional relation, $I = f(\lambda)$, and we plot the familiar curve as $1/I$ against λ . Actually we get two visibility curves, the photopic and the scotopic, which are similar in shape but which have their maxima at different wave-lengths.

The relationship represented by these curves is more complex than the simple function for density: $m = kV$. For a long time the function could be regarded merely as an empirical law measuring a derived magnitude and some other unknown quantities. More recently, however, these quantities have become known by a study of the nature of the ocular absorption of different wave-lengths, so that the law can now be formulated as a normal probability function.¹⁶ The function is: $v = e^{-0.0002705(\lambda - 554.9)^2}$, where v = visibility and e = logarithmic base. The constant, 554.9, becomes 511. in the case of scotopic vision. It is a measure of the maximal sensitivity of the eye, with reference to wave-length, under given conditions of light adaptation. This is an excellent example of the validation of a derived magnitude in research stimulated originally by the use of the magnitude!

It is interesting to note again, in this case, the effect of the 'focus of attention' on the locus of the real magnitude measured. When density is measured, we attribute the density to the object measured. When visibility is measured, we are inclined similarly to attribute the visibility to the different wave-lengths of light. However, the establishment of the different visibility curves for photopic and scotopic vision suggests at once that we are measuring here, not the visibility of the light, but the sensitivity of the photo-receptor processes in the retina. This shift of the magnitude measured, from the light to the retina, is typical of the way in which properties in the physicist's domain external to the organism may shift into the domain of the psychologist within the organism. In this connection it is interesting to note a recent suggestion that the Visibility Curve be renamed the Sensitivity Curve.

¹⁶ L. T. Troland, *Handbook of general experimental psychology*, 1934, p. 668.

PSYCHOLOGY AND PHYSICS

This interpretation of measurement has produced no valid distinction between psychological and physical magnitudes. It makes psychology and physics indistinguishable. So far as the nature of the magnitudes measured is concerned, such a conclusion is right and proper. There is no operational distinction, for example, between the measurement of density and the measurement of the visibility function. Length is as psychological as it is physical. Hardness is as mental as pain. Observation and its operational tool, discrimination, are the fundamental factors in all scientific measurement.

From this point of view we find that psychological measurement and physical measurement are one and the same thing. We can measure the 'mind' with as much logical validity as the physicist can measure uniform acceleration, because 'mental' magnitudes and 'physical' magnitudes are equally 'psychological' or equally 'physical.'

There is no doubt, however, that psychological and physical experimentation are directed toward different goals. The psychologist is interested in the explanation of the *modus operandi* of the human or animal organism. The physicist desires to explain the nature of phenomena within the realm of inorganic matter. The physicist is concerned about length because it enables him to learn more about the nature of relationships between the properties of inorganic matter. The psychologist studies length only insofar as a knowledge of it enables him to learn more about the nature of the organism. Though the physicist rejects the measurement of illusory visual length because it tells him little about the nature of light, the psychologist gladly accepts this same form of measurement because it offers data which may become invaluable in the construction of a theory of vision. Thus psychology steps boldly into the ranks of the natural sciences, using physical magnitudes and physical operations to explain a set of phenomena peculiarly its own.

Psychological measurement, understood in operational terms, is a *fait accompli*. It is physical measurement. It

always has been. And the psychologist, now aware that he is using no mysteriously unique scientific instrument (the observer), can, secure in his new self-knowledge, proceed with his measurements, unimpeded by the hampering difficulties of the Cartesian dichotomy between mind and body.

[MS. received October 31, 1934]

ON THE TACTILE PERCEPTION OF VIBRATION FREQUENCIES

BY WALTHER JOËL

University of Southern California

I

In the tactile perception of vibrations the physical stimulus is the vibration of a body in simple harmonic motion. This vibration is commonly analyzed into amplitude (a) and frequency (n). A third factor, often neglected in psychophysical studies, is the energy (E) of the moving system which practically equals the energy required to move the system. The formula

$$E = ka^2n^2 \text{ (where } k \text{ is constant)}$$

means, then, that the energy represented by a system moving in simple harmonic motion is proportional to both frequency and amplitude.

The experience produced by a vibrating body in contact with the skin is usually described as vibratory sensation. It may be more or less intense, *i.e.*, *strong* or *weak*, *hard* or *soft*. It may also vary in its quality. The terms most frequently used relative to quality are: *smooth-rough*, *fine-coarse*, *fast-slow*, *high-low*.

In the psychophysical study of vibratory sensations the subjects usually apply the terms *faster* and *slower* to describe the perceived difference between two vibrations successively presented as tactile stimuli. The assumption is then made that the perception expressed in these terms is based on the frequency of the vibratory stimulus provided that the amplitude, the supposed correlate of intensity, is kept constant.

Such a procedure contains two fallacies. First, it assumes that the subject's judgment of *faster-slower* is based on a variation in frequency only. And, second, it assumes that the terms *faster-slower* really describe the sensation.

That the first assumption is wrong should be easy to prove, since the essential requirement of any experiment—to keep all variable factors constant but one—is not met in this procedure. More energy is required to move a given mass four hundred fifty times per second through a certain distance (amplitude) than to move that same mass through the same distance only four hundred times per second. Therefore, *ceteris paribus*, the amplitude must increase with a decrease in frequency.¹ In other words, even with the same amount of energy, we cannot produce vibrations of different frequencies that will be equal in amplitude. One might make equal the amplitudes of two vibrations differing in frequency. But this could be done only by changing the energy which produces the vibration. A glance at the above formula shows that it is futile to attempt the control of all variables by keeping the amplitude constant. When two different frequencies with equal energy are compared, the slower frequency has of necessity the greater amplitude. When the amplitudes are made equal, the energies differ. It is, therefore, impossible to present two stimuli which differ in nothing but frequency. Hence two vibrations cannot be differentiated on the basis of frequency alone.

The second fallacy in assuming that the subject's judgment is based on frequency lies in the allegation that the terms *faster-slower* really describe the *felt* difference. But *faster* and *slower* are merely words indicating a discrimination of something. Even if we could present two vibrations differing in frequency, but otherwise equal—which, according to the above discussion, is impossible—the words *faster* and *slower* would not prove that the sensation is actually one of speed.

To avoid this difficulty, we might instruct the subject to report in other terms, *e.g.*, *smoother-rougher*, *finer-coarser*, or *higher-lower*. Then, as long as his reports vary consistently in the same direction as the vibratory stimulus, we may speak of a genuine discrimination. But what does he discriminate?

¹ As far back as 1892 Sergi (6) pointed to the decreasing amplitude of vibration with increasing frequency as the cause of apparently decreasing sensibility of the skin to successive stimulation with increasing frequency (between 50 and 1,000 dv.).

In the physical stimulus there are differences both of frequency and of energy. Which of these is the basis of discrimination?

II

This question was experimentally studied by the writer.² The stimulus pairs were 200-225, 400-400, 400-425 and 425-450 dv. The subjects in all experiments had to judge the second of two successively presented vibrations in comparison with the first.

a. Five subjects judged a total of 2,740 such stimulus pairs. In a portion of the trials the energy of both vibrations was kept so that the subjects could not detect an intensity difference. In the remainder the intensity was noticeably different. Half of the pairs were judged as to quality in terms of *smoother-rougher* or *finer-coarser*, the other half as to intensity in terms of *stronger-weaker*. There was no quality discrimination where intensity discrimination was impossible. But there was quality discrimination where intensity discrimination was possible.

b. Twenty subjects each judged one hundred stimulus pairs; fifty as to intensity, and fifty as to quality. In each fifty trials, twenty pairs were presented which could not be discriminated as to intensity, and thirty pairs with a perceptible intensity difference. In 89 ± 2.11 per cent of the cases, quality differences were perceived only where intensity differences were perceived too.

c. In another experiment, three subjects had to choose in each trial between quality and intensity judgment. Despite training, they did not learn to distinguish between frequency and energy differences of vibration.

The results of this study showed that the subjects discriminated, in terms of quality, vibrations of different frequencies only under such energy conditions that permitted them to discriminate also in terms of intensity. They even perceived a 'frequency' difference between two vibrations of the same frequency (400-400 dv.) when there was a discriminable energy difference present.

² W. Joël, The tactile perception of vibration frequencies, Doctoral Dissertation, University of Southern California, Los Angeles, 1934, 30 pp.

III

In the light of these findings it is necessary to mention briefly two other studies on the tactile discrimination of vibration frequencies.

Knudsen (3) made some exploratory trials on the frequency difference threshold for various frequency levels. What was the basis of discrimination? How did he solve the problem of intensity control?

The amplitude of all test frequencies was adjusted to fifty times the minimal threshold amplitude. This insured equality of intensity, on a sensation scale, of the two vibrations which were to be compared with regard to frequency of vibration. It also insured uniformity of intensity of all vibrations used for these tests (p. 335).

This method of equating intensity is not satisfactory. And further doubt is cast upon the genuineness of frequency discrimination by the following statement of Knudsen:

Before each test the observer felt the two vibrations which were to be compared, in order that he might recognize the distinguishing characteristics of different rates of vibration. After a little experience it was fairly easy to distinguish a difference of frequency provided the difference was large enough (p. 334).

This means, if we understand correctly, that the observers were trained to differentiate supposedly the frequencies. But what they actually judged is not certain, since we have no indication whether the subjective intensity was equal for the two vibrations.

Roberts (5) attempted to solve the intensity difficulty by comparing a vibration of standard frequency and amplitude with a vibration of different frequency and variable amplitude. He found that increasing the amplitude of the higher frequency rendered it less discriminable, until finally the judgments were reversed, *i.e.*, the higher frequency was judged lower. But there was clear discrimination where the amplitudes of standard and comparison vibrations were equal. Thus Roberts quotes as evidence

that a genuine frequency discrimination, in the strictest possible sense of the term, was obtained. . . . If we were dealing only with

an amplitude discrimination, confusion would occur close to the point of objective equality, *i.e.*, in the vicinity of one hundred per cent on the graph (p. 305).

But, as we have shown, the energy for vibrations of different frequencies and equal amplitudes is not equal. Therefore, the critical point is not where the amplitude of the comparison frequency is one hundred per cent of that of the standard. For at this point the judgment is not based on frequency difference alone. The point of confusion must rather be found where the amplitudes are not equal, and where the total impression resulting from unequal amplitudes, unequal frequencies, and unequal energies, is subjectively equal. Roberts' data actually bring out this fact.

Roberts spent more effort in training his subjects than Knudsen did. But training does not necessarily mean that the subject learns to *dissociate* various aspects of the stimulus and then to *associate* them with the adequate word. Rather he learns to associate a certain stimulus pattern with a word. Thus he will be able to judge habitually *faster* whenever a certain total stimulus pattern acts upon his tactile receptors. Energy, amplitude, and frequency are the important factors in this pattern. After the verbal response *faster* has become well associated with the total stimulus pattern, the subject is said to have learned to discriminate *frequency*. The fallacy is evident.

IV

We may now return to our original question: When vibrations of different frequencies are perceived as *rougher* or *smoother*, *slower* or *faster*, etc., what is the stimulus basis for this discrimination?

It is not the *amplitude* alone, since discrimination may take place with equal amplitudes.

It is not the *energy* alone, since discrimination may take place with equal energies.

It is not *frequency* alone, since it is impossible to produce two different vibrations that differ in nothing but frequency.

What, then, is the stimulus basis for the discrimination?

It must be frequency, amplitude, and energy. All three of them together form the stimulus basis, of which *intensity* is the subjective correlate as much as *quality*.

V

It is interesting, in this connection, to note the following parallel findings in the field of audition.

Knudsen (2) found that minimal changes in frequency and minimal changes in r.m.s. pressure may be discriminated as a change, but only as a change, so that the subject was not able to say consistently which was a change of pitch, which was a change of loudness.

Metfessel (4) reports that his subjects were unable to differentiate between a *pitch vibrato* and an *intensity vibrato*. He thinks they committed the stimulus-error in which their knowledge of the nature of the physical vibrato determined their reports. But the stimulus-error is being committed whenever we describe an experience. For by describing it we attach some meaning to it. Thus, the description of vibrato, a new type of experience, can only be made in terms of something known. Speaking, then, of pitch fluctuation presupposes the knowledge of pitch and that pitch can fluctuate. To ascribe *this* meaning to the new experience of vibrato is just as much an interpretation in terms of what we know about auditory stimuli as it would be to call it intensity fluctuation.

Fletcher (1) reports that "An observer's location of the pitch was found to depend upon the intensity as well as the frequency." This influence of intensity upon pitch perception was more marked in some parts of the frequency range than in others.

Very recently Stevens (7) has found that

pitch and loudness are different functions of both frequency and energy. . . . Two tones of different frequency can be made to sound equal in pitch or loudness by a proper adjustment of the energy of one of the tones. . . .

SUMMARY

1. It is impossible to produce two vibrations that differ in nothing but frequency.
2. Amplitude is not the stimulus correlate of tactile intensity when vibrations of different frequencies are concerned.
3. Tactile discrimination of 'frequency' made in such terms as *faster-slower*, *finer-coarser*, *smoother-rougher*, has not only frequency differences as its stimulus basis.
4. The experience of intensity depends upon frequency, amplitude and energy of the vibration.
5. The experience of quality, such as *roughness-smoothness*, depends upon frequency, amplitude and energy of the vibration.
6. Where, in other investigations, success in training to discriminate 'frequency' by touch has been claimed, the verbal response was apparently connected with the total stimulus pattern of vibration (including those factors usually considered as a basis for 'intensity' discrimination) and then mistaken for genuine frequency discrimination.
7. These considerations on the tactile perception of vibrations apply, under certain conditions, to hearing as well.

BIBLIOGRAPHY

1. FLETCHER, H., Loudness and pitch of musical tones and their relation to the intensity and frequency, *Science*, 1934, 79, p. 484.
2. KNUDSEN, V. O., The sensibility of the ear to small changes of intensity and frequency, *Phys. Rev.*, 1923, 21, pp. 84-102.
3. KNUDSEN, V. O., 'Hearing' with the sense of touch, *J. General Psychol.*, 1928, 1, pp. 320-352.
4. METFESSEL, M., The vibrato in artistic voices, from University of Iowa studies in the psychology of music, 1932, 1, pp. 14-117.
5. ROBERTS, W. H., A two-dimensional analysis of the discrimination of differences in the frequency of vibrations by means of the sense of touch, *J. Franklin Inst.*, 1932, 213, pp. 283-311.
6. SERGI, G., Über einige Eigentümlichkeiten des Tastsinns, *Zsch. f. Psychol.*, 1892, 3, pp. 175-184.
7. STEVENS, S. S., Pitch, loudness, volume and density, Program, Forty-second Annual Meeting, American Psychol. Assoc., 1934, p. 18.

[MS. received November 2, 1934]

THE EXPERIMENTAL APPROACH TO CONCEPT LEARNING

BY KENNETH L. SMOKE

Mary Baldwin College

Many of our difficulties in psychology arise from the fact that our terms are not defined clearly and unambiguously. 'Concept' is a case in point. Partly because of this failure to evolve a clear-cut meaning of 'concept,' and partly because many writers have viewed concepts solely in terms of the tenets of some 'school' of psychology, there is today a vast confusion in psychological literature concerning what it means to learn a concept.

From the systematic point of view there are at least as many theories concerning concepts as there are 'schools' of psychology.¹ It is not our purpose, however, to examine these theories as such. Hence we are not interested in either attacking or defending any 'school.' What seems to us to be important is that those who claim to be dealing experimentally with concepts shall all be talking about the same thing. Thus it is our purpose to attempt to formulate a way of approaching concept learning that can serve as a basis for systematic experimentation without necessitating adherence to any 'school.' We shall approach our task by way of a brief survey of the meager experimental literature on concept learning.

Prior to 1920 all of the experimental work on concepts was carried on by means of an introspective technique. The studies by Moore (6) and Fisher (3) are cases in point. This

¹ A comparison of the following quotations will illustrate the way in which differences that exist among the 'schools' are reflected in definitions of 'concept': "A concept . . . is a group of responses which have been classified together and labeled by a verbal symbol" (J. Stanley Gray, A behavioristic interpretation of concept formation, *PSYCHOL. REV.*, 1931, 38, 72); "The concept is an image functioning in such a way as to suggest a definite meaning, or core of meanings, which the mind attaches equally to all the individuals of a group, or species" (I. E. Miller, *The psychology of thinking*, New York, Macmillan, 1917, 192).

work, though careful and painstaking, is so lacking in objectivity (in the sense in which the natural sciences are objective) that it seems to us to be of little value today.

The first important non-introspective attempt to study concept learning was that of Hull (5). We have already published a critique of this work (9, 2-5). Here it is sufficient to say that Hull considered his subjects as having 'evolved' a concept when they discovered the 'common element' hidden in each member of a group of Chinese characters. Each character in a given group had this 'common element' in the sense that it contained 'certain strokes in common' with the other characters in its group, and the process of discovering these 'common elements' was taken to be 'the evolution of concepts.'² Can the reader think of a single concept that he learned in this manner?

In more recent years Ewert and Lambert (1) have made an entirely different experimental approach to concept learning. Using J. C. Peterson's problem in which the learner tries to transfer a number of discs from one circle to another by way of a third, these investigators attempted to study 'the effect of verbal instructions upon the formation of a concept.' "The elements of the problem remain the same, no matter how many discs are used, so that it is possible to make a generalization or to form a concept of the solution for any number of discs" (1, 402). Here the 'concept' that is being investigated is a formula or principle. The phrase 'the formation of a concept' in the title of the study under consideration clearly means the discovery of a formula or principle. Apparently in much the same vein, Murphy, speaking of a study by J. C. Peterson (8) in which the learner tried to discover the rule by which a certain game can be won, calls this study 'an experiment on the formation of concepts' (7, 388). The identifica-

² The phrase 'the evolution of concepts' is very appropriate, as anyone who has given an intelligence test to three-year-olds and to twelve-year-olds can testify. At the beginning the learner, especially the youthful learner, is likely to conceive of anything merely in terms of the use to which it can be put, or at least in terms of the use to which he has put it. Thus, the small child's answer to "What is a ball?" may be "A ball is to play with." The species-genus classification of a dictionary definition is likely to be the expression of a very sophisticated and highly evolved concept.

tion of concepts with such things as rules, formulae, and principles is logically fallacious and psychologically confusing. To be sure concepts are indispensable constituents of rules, formulae, and principles, but there is no justification for identifying concepts with these.

In another recent experimental approach to concept learning white rats were employed. Thus, Fields (2) has written a monograph bearing the title, 'The development of the concept of triangularity by the white rat.' In defense of his view that white rats have concepts, Fields writes,

If Hull is justified in using the term 'concept' to denote the observable behavior of a child when it makes a specific language response to a certain stimulating situation, then we should be allowed to use the term 'concept' to denote the observable behavior of a rat when it makes a specific muscular response to a similar stimulating situation. The difference seems to be one of degree, not of kind, in that it is necessary for the rat to use less complex reaction units than the child. But so far as the present study is concerned, the specific acquired reaction involved in the 'jump to triangle' of the rat will be considered as equivalent to the child's specific language response 'triangle' (2, 4).

But is the "jump to triangle" of the rat . . . equivalent to the child's specific language response 'triangle'? Is there not a symbolic aspect of the child's language response 'triangle' which is not demonstrated by the mere 'jump to triangle' of the rat? The child's language response 'triangle' *stands for or is symbolic of* certain things. We are unable to see that there is any good evidence that the rat's response stands for or is symbolic of anything, although it clearly is discriminatory of certain things. Perhaps it would be safer to merely say, as Gengerelli does, that "the white rat is capable of certain generalized habits of a limited kind" (4, 200).

The non-introspective experimental approach to concept learning has thus been characterized by (1) the tendency to regard concept learning as a process of discovering a hidden 'common element,' (2) the tendency to regard concept learning as a process of finding the rule, formula, or principle applicable

in a given case, and (3) the tendency to regard concept learning as a process of discriminating such things as triangles. There appears to be no hope of reconciling such divergent views. Is there then any common ground with reference to concept learning upon which non-introspective experimenters can meet? We believe that there is, and offer our view of it.

"By 'concept formation,' 'generalization,' or 'concept learning,' we refer to the process whereby an organism develops a symbolic response³ (usually, but not necessarily, linguistic) which is made to the members of a class of stimulus patterns but not to other stimuli" (9, 8). Concept learning never occurs unless there is a response on the part of the learner to the relationships common to two or more stimulus patterns. Some of our more abstract concepts (such as 'honesty,' 'liberty,' 'patriotism') are perhaps learned from stimulus patterns that have *only* relationships in common. Other concepts (such as 'the Union Jack,' 'mountain,' 'skyscraper'⁴) are learned from stimulus patterns that have not only relationships but also certain characteristics (color, size, etc.) in common. But the *sine qua non* of concept learning is always response to relationships present in each member of a group of stimulus patterns, the stimulus patterns in question being classified as a group by virtue of the fact that they have certain relationships, and perhaps certain characteristics, in common.^{5, 6}

The experimental psychologist must have a standard in terms of which he can judge whether his subject has learned a concept. Present knowledge does not make possible the

³ We are convinced that the symbolic aspect of the response in the case of concepts has not received the attention which it deserves. This is true to some extent in our own work. Thus, instead of telling the subject that the name of a type of design is 'vec' and then, after the learning period, testing whether he has learned 'vec' by having him write 'Yes' or 'No' when discriminating 'vecs' from other designs, it would be better to require him to write 'vec' or 'not a vec' in response to each design used in testing. We are making this change in technique in some work now in progress.

⁴ These illustrations are not to be taken as implying that words and concepts are one and the same. A word is not a concept, but a concept always involves a word or some other instrument for carrying the symbolism.

⁵ This view of concept learning has already proved to be useful in a series of experiments. See (9) and (10).

⁶ As a possible aid in the study of the genesis of these classifications, we propose the concept of a 'conceptual DL.'

setting up of a biophysical criterion of concept learning, that is to say, it is at present impossible to indicate the neuromuscular and the neuroglandular events that must occur before an individual may be said to have learned a concept. It is entirely possible, however, to set up a biosocial criterion. We propose the following: *consistency of differential, generalized, symbolic response*. The standard by which one can judge whether a given individual has learned a given concept, we suggest, is this: the consistency with which he is able to make symbolic responses that differentiate the members of the class of stimulus patterns in question from stimuli which are not regarded as falling in that class. To "have a concept of ' ' " means to have such a neuromuscular and neuroglandular organization that one can consistently make symbolic responses that differentiate stimulus patterns which fulfill the conditions essential to ' ' from those that do not.

This way of envisaging concept learning seems to us to leave the door open to any non-introspective experimental approach to the problem, regardless of the systematic convictions entertained by the experimenter. Thus, although we have not spoken of 'awareness' or 'meaning,' we see no conflict between what has been said and the view maintained by some psychologists to the effect that such concepts as these are essential in a complete interpretation of many of the phenomena which psychologists study. It is our belief that the thing which we have attempted to do in the case of concept learning can be done in the case of all other problems in the field of psychology, that is to say, we believe that all psychological problems can be viewed in terms of stimulation and response for the purpose of experimental attack upon them without thereby destroying the validity of such concepts as 'awareness' and 'meaning.'

The problem of concept learning is a key problem in the field of the 'higher mental processes.' It is therefore not surprising that the confusion that characterizes the literature on concept learning is also present in the case of other problems in this field. The continual accumulation of studies in which the term 'concept' is used with reference to radically different

sorts of processes will merely serve to make our present confusion worse confounded. The foregoing is presented in the hope that it meets the need for a clear-cut way of viewing concepts and that it offers a theoretical basis upon which investigators can reach some degree of agreement in the experimental approach to concept learning.

REFERENCES

1. EWERT, P. H., AND LAMBERT, J. F., The effect of verbal instructions upon the formation of a concept, *J. General Psychol.*, 1932, 6, 400-413.
2. FIELDS, PAUL E., Studies in concept formation. I. The development of the concept of triangularity by the white rat, *Comp. Psychol. Monog.*, 1932, 9, No. 2.
3. FISHER, SARAH C., The process of generalizing abstraction; and its product, the general concept, *Psychol. Monog.*, 1916, 21, No. 2.
4. GENDERELLI, J. A., Studies in abstraction with the white rat, *Ped. Sem. and J. Genet. Psychol.*, 1930, 38, 171-202.
5. HULL, C. L., Quantitative aspects of the evolution of concepts, *Psychol. Monog.*, 1920, 28, No. 1.
6. MOORE, T. V., The process of abstraction, *Univ. of Calif. Publ. in Psychol.*, 1910, 1, No. 2.
7. MURPHY, GARDNER, General psychology, New York, 1933.
8. PETERSON, J. C., The higher mental processes in learning, *Psychol. Monog.*, 1920, 28, No. 7.
9. SMOKE, KENNETH L., An objective study of concept formation, *Psychol. Monog.*, 1932, 42, No. 4.
10. SMOKE, KENNETH L., Negative instances in concept learning, *J. Exper. Psychol.*, 1933, 16, 583-588.

[MS. received November 19, 1934]

A REPLY TO 'SIGN-GESTALT OR CONDITIONED REFLEX?'

BY NEAL E. MILLER

Institute of Human Relations, Yale University

In a very interesting and provocative paper, Professor Tolman (12) has maintained that, on the basis of sign-gestalt theory, 'insightful' behavior ought to occur in a certain learning situation. He has also expressed the opinion that, if this behavior actually does occur, it will be unexplainable in terms of conditioning.

Professor Tolman has devised an ingenious experiment to illustrate his point. Four rats were trained to run from a starting box down a short alley, make a black-white discrimination, enter a food box, and secure food. After they had learned this thoroughly, they were placed *directly* in the food box, without having run the alley, and given an electric shock there. Then each animal was given a test run down the alley. It was thought that the cues in the alley should now lead to expectations of shock instead of expectations of food so that the animals should refuse to run. Since the experience of running the alley had never immediately preceded the shock, it was believed that the refusal to run would be unexplainable in terms of conditioning. Thus the experiment was to be a crucial disproof of the adequacy of conditioned response theory.

However, it so happened that, to quote the paper (12), "Each rat after having been shocked in the food compartment and then carried to the starting point, immediately dashed off

¹ Read before the American Psychological Association, New York meeting, September, 1934. Acknowledgments are due to Professor Walter R. Miles for reading and criticizing the manuscript, to Professor Clark L. Hull for donating the rats used in the first and third experiments, especially to Miss Elizabeth A. Ferguson for doing, under the direction of the author, a major portion of the work on the first experiment, and to Dr. Florian Heiser for allowing her to do this work as a project for his course in Experimental Psychology.

gaily and just as usual through the whole discrimination-apparatus and bang whack into the very food compartment in which he had just been shocked." This outcome was not in line with the sign-gestalt prediction.

The apparent stupidity of the rats might seem to be wholly in line with the principles of conditioning. However, it is our opinion that such a view would be based on a misconception of the potentialities of conditioned-response mechanisms. It is true that the intelligent behavior expected by sign-gestalt would not itself be a simple conditioned reaction. But there is reason to believe that it should be produced, as something new, by the interaction of several familiar principles of conditioning. In fact, a conditioned response analysis of the sign-gestalt experiment reveals a probable reason for the negative results and leads one to predict that the insightful behavior should appear when the experiment is appropriately modified.

It should be understood that the following analysis is put forward only as a tentative hypothesis and that the validity of its conclusions will of necessity depend upon the validity of the assumptions involved. Having made a specific statement of these general reservations, we shall omit encumbering qualifying phrases from the remainder of the discussion.

The discrimination apparatus, which happened to be a part of Professor Tolman's set up, is not essential to the issue at stake. So, let us eliminate it and consider the case of a hungry animal running down a single alley for food. The animal already possesses responses to joint stimulation from hunger and food. Following Professor Hull's (2) terminology, we shall call these responses, such as eating, salivating and related behavior, the goal responses. Since the goal responses, reinforced by food, regularly take place immediately following and during stimulation from the cues in the food box, they should become conditioned to these cues.

But the reinforced goal responses also occur regularly a short interval after stimulation by the cues from the starting box and from the several portions of the alley. Now, it is well known that if reinforcement regularly occurs a short interval

after a given stimulus, a trace conditioned reaction to that stimulus tends to be set up and that there is a strong tendency for this trace conditioned reaction to antedate its reinforcement (7), (10). Thus there should be a strong tendency for the goal responses to occur in the alley and even in the starting box.

This theoretical expectation is given convincing empirical support by the well established fact that even reactions near the goal, such as the last turn leading to food, show a strong tendency to intrude into the earlier portions of a maze sequence (8), (9). However, it is practically impossible for the complete goal reaction to take place without the actual presence of food. So one would ordinarily expect only some parts of the goal reaction to take place in the starting box, the alley, and in the section of the end box preceding the food. These might be such items of behavior as incipient mouth and head movements. Let us follow Hull's (2) terminology again and call these antedating, fractional, components of the goal response which should occur in the starting box and the alley, *anticipatory goal responses*.²

Now, the principles of physiology lead us to expect these antedating responses to produce a certain amount of characteristic interoceptive stimulation.³ Thus, when the animal is placed directly into the food box and then shocked, the cues there will necessarily be expected to arouse anticipatory (or complete) goal responses and the stimuli from these responses must precede and coincide with the shock. The responses to the shock will therefore be conditioned to these interoceptive stimuli.

When the hungry animal is placed in the starting box, the cues there will, as has been shown, lead to similar anticipatory goal responses. These responses will produce some of the

² Anticipatory is used in the objective sense and refers only to the position of the response in a temporal series.

Since short trace conditioned reactions are easier to establish than long ones (7), (14), the anticipatory responses nearer the goal should be stronger (at least during the earlier stages of training) than those further from the goal.

³ Interoceptive stimulation is meant to include all stimulation from sources within the organism: i.e., proprioceptive and enteroreceptive stimulation.

interoceptive stimuli to which the reactions to the shock have just been conditioned. The animal should, accordingly, be expected to exhibit in the starting box or the alley certain components of the withdrawing, crouching, sitting behavior which characterizes a rat's conduct following a shock.⁴ Thus, we have deduced from strict conditioned response principles that behavior related to the shock should appear in a situation (the starting box and alley) which, to state it anthropomorphically, might on the basis of past experience be expected to lead to the shock even though this situation had *never* itself directly preceded the shock. This type of transfer of training is commonly called foresight.

Let us restate our analysis in brief, dogmatic form. The laws of trace conditioning demand anticipatory goal responses. The fact that the same reinforcement follows both the food box and the alley requires that the anticipatory goal responses to these two be at least partially identical. The interoceptive stimulation from these similar anticipatory responses is the common element by which the conditioning of the shock is transferred from the food box to the alley.

But the sign-gestalt rats did not exhibit the foresightful behavior demanded by the conditioned response analysis. How can this be accounted for? It will be remembered that the foresightful behavior was deduced from the assumption that a distinctive response, affording a characteristic pattern of interoceptive stimulation, was present in the food box. If this response was absent during the original training to run the alley, components of it could not become anticipatory. If this response was absent or only weakly functional when the animals were placed directly into the food box and shocked, the reactions to the shock could not become conditioned to any characteristic pattern of interoceptive stimulation. In either

⁴ We say 'certain components' because this paper is an attempt to apply the general principles of conditioning rather than to break learned behavior up into any ultimate units. Other studies (4), (5), and (6) have supplied direct evidence that at least four of the major phenomena of conditioning experiments—experimental extinction, conditioned inhibition, spontaneous recovery, and a caffeine effect opposing experimental extinction—are general enough principles of behavior to be applicable to the situation of a hungry animal running down an alley for food.

case the insightful behavior would not manifest itself.⁵ If, on the other hand, an experiment should be arranged in which the distinctive response affording the crucial interoceptive stimulation was sure to be present in a gross form, the insightful behavior should appear.

To test this logic, several experiments were set up in which the reward devices illustrated in Fig. 1 forced the animals to

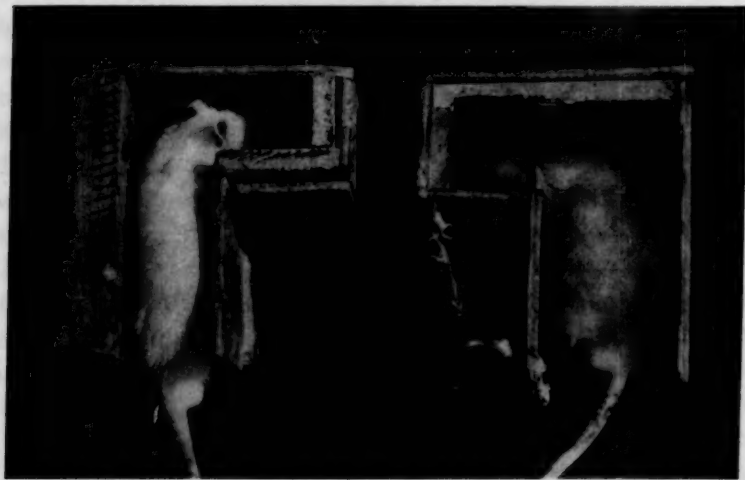


FIG. 1. The reward devices requiring the rats to make distinctive responses.
(Viewed from above.)

make very gross and distinctive responses in order to secure the lure after entering the end box of the alley. In one reward device the animal was forced to climb up and make a sharp turn to the right to secure food; in the other he must go straight in and make a sharp turn to the left to secure water.

In the first experiment seven animals, motivated by both hunger and thirst, were trained to run down a short alley at the end of which they found only the special reward device containing food. At other intervals during this preliminary learning, these animals were placed in a special cage, isolated

⁵ In the interests of a simple, clear exposition, we are not elaborating another possibility, less charitable to the rat, namely, that these animals are incapable of acquiring distinctive anticipatory reactions under any conditions. Subsequent results do not favor this hypothesis.

from the alley, and given separate training to use both the food and the water reward devices there, one at a time. The whole set-up is illustrated in Fig. 2.

Then three of the seven animals were put through the following procedure: The rat was placed in the special isolated cage which contained only the food in its reward device. As soon as he had climbed up, turned to the right in the accus-

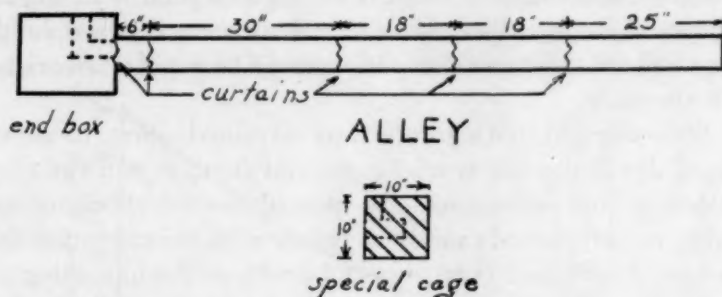


FIG. 2. Diagram of the experimental set-up. The dotted lines in the end-box and in the special cage indicate the position in which the food reward device was placed during the trials involving its use. The water reward device was used in a symmetrical position on the opposite side.

tomed manner and taken his first bite of food, he was shocked for one second.⁶ Then he was given time to get off the grid and was kept in this special cage for a two minute period during which he was given an additional, immediate shock each time he touched the grid. At the end of this period he was taken out and placed in another cage for three minutes in order to allow the general disturbance from the shock, which Pavlov would call external inhibition, to subside somewhat. After this, he was placed in the starting box of the alley. The length of time it took him to reach the last curtain (illustrated in Fig 2) was recorded. The reward device was now absent from the box at the end of the alley. Thus, the animal could be put back into the starting box and given a total of five test runs without danger of secondary conditioning.

⁶ The current for the shock was delivered from an inductorium through a 50,000 ohm resistance to a grid of wires stretched tightly across an air gap. While receiving the shock, most of the animals hunched themselves together and jumped up and down, not leaving the grid till the current ceased.

The other four rats were put through a similar procedure except that they were shocked in the isolated special cage while using the water reward device.

The experimental results confirmed the deduction from the conditioned response theory. The animals that had been shocked in the isolated food device while performing the act identical with the one ordinarily following their alley behavior sequence, required longer on the average for their first test run through the alley than those that had been shocked in the water reward device while performing the act *not* associated with the alley.

After this test, both groups were retrained—first, to use the reward devices in the special cage, and then, to run the alley. Following this retraining, the procedures of shocking and testing were repeated exactly as before with the exception that the rats, which had been shocked previously while using the water reward device, were shocked now while using the food reward device, and vice versa.

The results of this retest lent additional confirmation to the stimulus-response prediction. The animals shocked in the food device were again slower than those shocked in the water device. Furthermore, a comparison could now be made between the effects of the two procedures on each rat. In this comparison it was found that six out of the seven animals took longer to run the alley on the first trial after the shock in the isolated food device than they did after the shock in the isolated water device.

In order to eliminate any constant errors possibly arising from differences either in the motivation with which the animals entered the two reward devices or in the shock delivered by each of them, sixteen new rats were used in a second experiment. Whereas, in the first experiment all of the animals were trained to run the alley for food, in this experiment eight of the animals were trained to run the alley for food and eight for water. In other respects the procedures in the two experiments were identical.

In the training trials of the second experiment, it was observed that the animals that were running the alley on their

way to enter the food reward device, which was placed at the right side of the end box and in which they were required to make a right turn, tended to run down the right side of the alley (See Fig. 2). On the other hand, those animals that were on their way to enter the water reward device, which was placed on the left and in which they must make a left turn, tended to run down the left side of the alley. The percentage of animals passing under each successive curtain off to the side, right or left, corresponding to their ultimate goal response is indicated in Fig. 3. These data are for the day preceding the

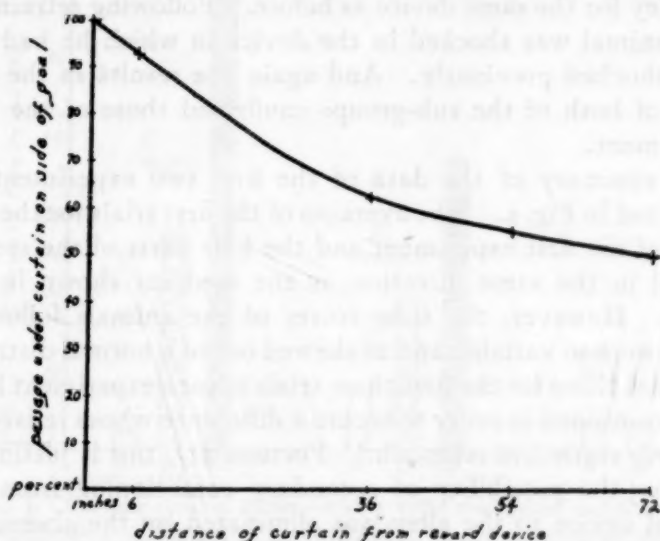


FIG. 3. Anticipatory responses in the alley on the trials immediately preceding the test runs.

test run. Thus, we see that the existence of the anticipatory goal response, stronger nearer the goal, which was postulated from theoretical considerations, can be verified by direct observation.

In the test trials of this experiment, half of the group running for food and half of the group running for water were shocked in the isolated special cage while using the device which each had found in the end box of the alley. The other half of each group was shocked in the isolated cage while using

the other reward device. The results in the test trials of both of the above sub-groups confirmed those of the first experiment. In other words, the animals which were shocked in the isolated reward device that had been associated with the alley—irrespective of whether it was the food or water device—took more time to run the alley in the first test trial than did the animals which were shocked in the device that had not been associated with the alley.

After the tests, the animals were retrained—first, to use each of the reward devices in the special cage, and then, to run the alley for the same device as before. Following retraining, each animal was shocked in the device in which he had not been shocked previously. And again the results in the test trials of both of the sub-groups confirmed those of the first experiment.

A summary of the data of the first two experiments is presented in Fig. 4. The averages of the first trials for the two parts of the first experiment and the four parts of the second are all in the same direction as the medians shown in the figure. However, the time scores of the animals following shock were so variable and so skewed out of a normal distribution that those for the first three trials in each experiment have to be combined in order to secure a difference whose reliability is above statistical reproach.⁷ Fortunately, this is justifiable because the possibility of secondary conditioning from the reward device to the alley was eliminated by the absence of this device from the end box during these trials. When these trials are combined, it is found that seventy-eight per cent (standard error .087) of the animals took longer to run the alley after being shocked outside it during the act identical with the one ordinarily following their alley sequence than they did after being shocked during the other, directly comparable act not associated with the alley. This is a difference

⁷ This variability and relative weakness of the foresightful behavior is exactly what would be expected from the nature of the mechanism postulated by the theory. If the anticipatory goal responses were absent or only weakly present either just before the shock or in the alley, the reaction to the former would ordinarily be weaker. Even with human beings the foresightful responses supposedly aroused indirectly are often over-ridden by more direct responses.

3.2 times its standard error above chance expectation and in the direction demanded by theory.

The fact that the acts performed in the two reward devices contain certain common elements, such as walking on the

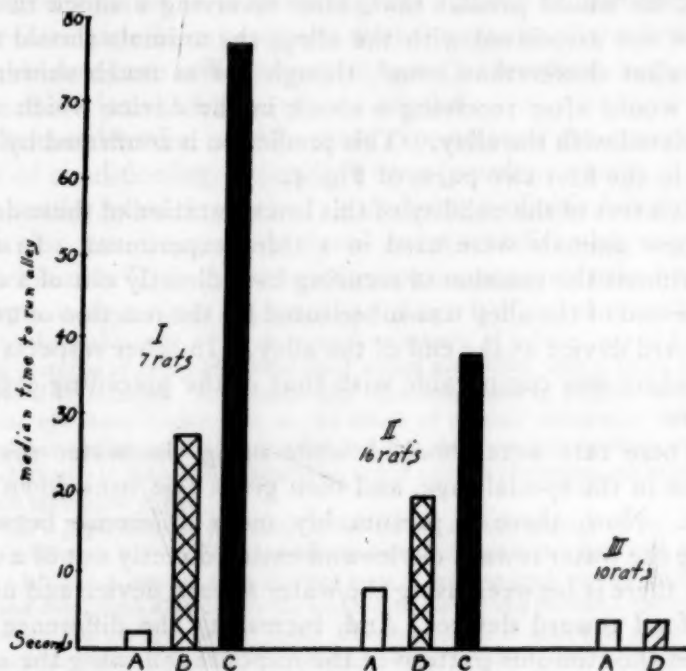


FIG. 4. The time required to run the alley under the various conditions of the three experiments. *A*, the time on the first trial of the day preceding the test runs; *B*, the time on the first test run after a shock in the isolated reward device which had not been associated with the alley; *C*, the time on the first test run after a shock in the isolated reward device which previously had been associated with the alley; *D*, the time on the first test run after a shock exactly like the one used in "*B*" but with the response which had followed the alley sequence during the preceding training different so that there was less similarity between it and the one preceding the shock.

grid, leads to the theoretical expectation that there should be a transfer of conditioning from the one to the other.⁸ This expectation is confirmed by the finding that, after a shock in one of the devices, all but four of the sixteen animals in the second experiment showed very noticeable tendencies to avoid

⁸ Generalization by similarity. For studies of this type of generalization see Bass and Hull (1) and their references.

the other device. This means that the effects of the shock administered in the device not associated with the alley must, in the preceding experiments, have been transferred to some extent to the device which was associated with the alley. Thus, we would predict that, after receiving a shock in the device not associated with the alley, the animals should run somewhat slower than usual, though not as much slower as they would after receiving a shock in the device which was associated with the alley. This prediction is confirmed by the data in the first two parts of Fig. 4.

As a test of the validity of this interpretation of these data, ten new animals were used in a third experiment. In this experiment the reaction of securing food directly out of a dish at the end of the alley was substituted for the reaction of using a reward device at the end of the alley. In other respects the procedure was comparable with that of the preceding experiments.

These rats were shocked while using the water reward device in the special cage, and then given test runs down the alley. Now, there is presumably more difference between using the water reward device and eating directly out of a dish than there is between using the water reward device and using the food reward device. And, increasing the difference between the stimulus pattern of the response following the alley and that of the response preceding the shock should lessen the transfer of conditioning from the one to the other. Thus, according to the hypothesis being tested, the animals in this experiment should not be slowed down as much by the shock in the device not associated with the alley as were the animals in the first two experiments. And, this is exactly what is shown by the medians in the third diagram of Fig. 4. Unfortunately, however, the one rat that had been variable throughout training refused to run at all after the shock. This necessitated the use of medians and spoiled the statistical reliability of the outcome.

In conclusion, the results of the first two experiments demonstrate that there is a generalization of conditioning (in other words, a transfer of training) from the reward device to

the alley, which is dependent upon the previous association between the two.⁹ A comparison of these results with the negative results of the sign-gestalt experiment, (12), suggests that the distinctiveness of the goal (or near goal) reaction is an important factor in determining the amount of this transfer. The third experiment suggests that the results of the first two would have been still more striking if the reactions in the two reward devices had been more dissimilar.

In the situation used in these experiments, the generalization of conditioning, dependent upon previous association, is commonly called foresight. A deduction of this foresightful behavior has been made from strict conditioned response principles.¹⁰ This conditioned response analysis may be said

⁹ This tends to confirm a hypothesis, put forward in a previous paper (3), that functional stimulus units, acting as channels for transfer of training, may be built up by the operation of an environmental trend toward consistency upon a native level of behavioral organization. It is hoped that experiments now in progress on human subjects will throw further light on the effects of previous associations upon the generalization of subsequent conditioning and that the theoretical significance of this type of transfer in the realms of perception and of neurotic phenomena can be pointed out in a subsequent paper.

¹⁰ In a paper recently presented before the New York meeting of the A.P.A., Professor Tolman has described another similar experiment (13). Rats were trained to turn left at a T junction of an elevated maze in order to reach food by the shorter path and then placed directly into the food box on the left and shocked. On the next run twenty-five per cent turned right. The assumption of somewhat different responses in the boxes to the left and to the right will adapt the analysis in this paper to that situation. When the animal is wavering, making tentative starts first in one direction and then in the other as these animals did, a turn to the left will confront the animal with cues arousing anticipatory responses creating stimuli to which have been conditioned reactions conflicting with running; a turn to the right will have less of a tendency to produce such reactions and so will be more likely to be continued. (Eventually turning reactions might become conditioned to stimuli created by conflict thus bringing about a more decisive response even in new situations.) Blind animals failed to show a transfer in this experiment (13), presumably because some of the stimuli leading to distinctive responses in the two food boxes were absent for these animals. According to our analysis, which does not yet pretend to be nearly complete, the blind animals should exhibit foresightful behavior if they are given specific non-visual stimuli leading to extensive reactions which are very different in the two food boxes.

The same type of deduction can be adapted to the 'Insight-Maze' described as the second experiment in Professor Tolman's first paper (12). Here the assumption has to be made that there is a response in the common path shared by the two shorter paths which is somewhat different from the responses occurring in the longest path, that the inhibitions (or conflicting responses) aroused by the barrier are conditioned

to have reduced the concepts used in the sign-gestalt description of foresightful behavior to more basic ones which are amenable to direct investigation and which already have been found useful in the description of hosts of other, quite different situations. The theoretical analysis using these stimulus-response concepts tends to be confirmed by the fact that it suggested the modification which produced the positive experimental results.

REFERENCES

1. BASS, M. J., AND HULL, C. L., The irradiation of a tactile conditioned reflex in man, *J. Comp. Psychol.*, 1934, 17, 47-65.
2. HULL, C. L., Goal attraction and directing ideas conceived as habit phenomena, *Psychol. Rev.*, 1931, 38, 487-506.
3. MILLER, N. E., The perception of children: a genetic study employing the critical choice delayed reaction, *J. Genet. Psychol.*, 1934, 44, 321-339.
4. MILLER, N. E., AND MILES, W. R., Alcohol and removal of reward: an analytical study of rodent maze behavior, *J. Comp. Psychol.* (in press).
5. —, Effects of caffeine on the running speed of hungry, satiated, and frustrated rats, (to be published).
6. MILLER, N. E., AND STEVENSON, S. S., Agitated behavior of rats during experimental extinction and a curve of spontaneous recovery, (to be published).
7. PAVLOV, I. P., Conditioned reflexes, Oxford University Press, 1927.
8. SPENCE, K. W., AND SHIPLEY, W. C., The factors determining the difficulty of blind alleys in maze learning by the white rat, *J. Comp. Psychol.*, 1934, 17, 423-436.
9. SPRAGG, S. D. S., Anticipatory responses in the maze, *J. Comp. Psychol.*, 1934, 18, 51-73.
10. SWITZER, S. A., Anticipatory and inhibitory characteristics of delayed conditioned reactions, *J. Exper. Psychol.*, 1934, 17, 603-620.
11. TOLMAN, E. C., Purposeful behavior in animals and men, New York and London: The Century Co., 1932.
12. —, Sign-gestalt or conditioned reflex?, *Psychol. Rev.*, 1933, 40, 246-255.
13. —, A case of learning in rats not explainable by conditioning. Read before the American Psychological Association, New York meeting, September, 1934.
14. WOLFLE, H. M., Conditioning as a function of the interval between the conditioned and the original stimulus, *J. General Psychol.*, 1932, 7, 80-103.

to the stimuli produced by this response, and that, as is well known (8), (9), responses some distance from the goal are also brought forward. This experiment failed to produce the results expected by sign-gestalt when tried in an alley maze (11, p. 167). The stimulus-response analysis demands that it could be made to work if the animals were forced to perform some very distinctive reaction, such as climbing over a hurdle, in the common path just before the place where the block was to be inserted.

[MS. received October 15, 1934]

THE TWO-FACTOR THEORY AND ITS CRITICISM

BY HENRY E. GARRETT

Columbia University

In a recent issue of the REVIEW (7), Professor Spearman has answered briefly certain criticisms of his Two-Factor theory made by Professor R. C. Tryon (8, 9). For psychologists familiar with factor theories, Spearman's reply will doubtless be entirely adequate, and further comment upon this controversy unnecessary. Because of the length and vehemence of Tryon's original statements, however, and his rather tart answer to Spearman's comments (10), I suspect that many psychologists working in other fields may reach the conclusion that Tryon has demolished the Two-Factor theory, or else shaken it badly. And this seems all the more likely since even competent workers in the field of mental organization seem to be impressed. Flanagan, for instance, in his recent excellent study of personality (3) speaks of Tryon's 'severe criticism' of Spearman's techniques; and of his having 'cogently demonstrated' that the tetrad criterion is not satisfied by the correlational data obtained from mental tests.

It would be unfortunate, I think, if an impression of this sort became general. Once an erroneous idea gets into the text-books, it is passed on by succeeding writers and becomes increasingly difficult to eradicate. I believe it is worthwhile, therefore, to point out in some detail the fallacious nature of the evidence upon which Tryon bases his attack upon the Two-Factor theory. I am not primarily concerned in this paper in championing any particular theory of mental organization; but I shall try to show that either a relatively few broad group factors or Spearman's g and s offer a better theoretical explanation of individual differences than do numerous small specific factors. Incidentally, I shall hope also to clarify several points bearing upon the interpretation of

factors, upon which there seems to be considerable misunderstanding.

1. EXPERIMENTAL EVIDENCE ALLEGED TO DISCREDIT THE TWO-FACTOR HYPOTHESIS

In his first paper, Tryon (8) assembles ten researches in which tetrad analysis of test data was carried out. These include an older study by Bonser, one study by Kelley, three studies by Spearman and his associates, three studies from the Columbia laboratory, part of the data from the Minnesota study of mechanical ability, and selected data from the character investigations of Hartshorne and May. For each study a comparison was made of the distribution of tetrads to be expected if only two factors (g and s) were present, and of the observed distribution of tetrads—those actually obtained. While no statistical test was made of the reliability of the differences between theoretical expectation and actual observation, an inspection of the graphs offers little evidence of close agreement between theory and observation. Tryon concludes from this fact that "there is not a shred (of evidence) which supports the view that such factors (g and s) are at work" (p. 350). This statement would be more accurate if changed to read, "there is no evidence that *only* g and s are at work." But in either form it has no significance as far as the truth or falsity of the Two-Factor theory is concerned. Tryon fails to point out the very important fact that none of the investigators cited made the claim that his data, *as it stood*, supported the Two-Factor hypothesis. Multiple factors were not only postulated, but in most cases carefully measured. Kelley, for example, calculated six factors to explain his correlations, while Stephenson was concerned with verbal *and* non-verbal abilities, *i.e.*, directly with group factors. Schneek's and Anastasi's experiments were planned specifically to measure group factors, and not g and s , so that the satisfaction of the tetrad criterion for *all* of their data was highly improbable. It is scarcely debatable that one cannot try out a theory wherever he likes, and then if it doesn't work triumphantly proclaim that it is false.

In his second paper, Tryon (9), recognizes at last the hitherto ignored group factors. By means of an ingenious scheme for locating the probable positions of 'bonds' (which is the major, if not the only contribution of both papers), he attempts to explain away the group factors postulated by the various authors already cited. Criticism here seems to be directed mainly against the *names* given by these writers to their group factors. It is not especially pertinent to any theory of mental organization, nor, as far as I can see, to the interpretation of the various factors themselves. An investigator has the right to call his group factors 'similarity of material,' 'mental speed,' 'ability to employ spatial relations,' or what-not so long as each factor is related directly to the group of tests in which it appears. The criteria set up for a group factor are the same common sense considerations used when we call a test one of memory, or perception, or introversion. Validity in such cases is largely a matter of definition.

2. IMPLICATIONS OF THE SAMPLING (MULTIPLE FACTOR) THEORY WHEN THE NUMBER OF FACTORS IS VERY LARGE

In proposing to substitute Thomson's sampling theory for the Two-Factor theory, Tryon emphasizes again and again the need for very numerous independent factors or element in order to explain individual differences. Spearman (6) has traced out the implications of the assumption that behavior is governed by the chance operation of a large number of equal and independent elements (as in dice throwing); and has concluded that "a large number of random factors will reduce the scatter of the individual differences toward zero." According to Spearman, therefore, the multiple factor theory (at least in its extreme form) is untenable, since the operation of a large number of chance factors would tend to give every one the same total ability. Tryon summarily (8) dismisses this argument as follows: "Spearman has repeatedly stated that under the independent multiple factor theory, a large number of factors would lead to all individuals obtaining the same total score. This proposition may be tested by dice throws, and amounts to saying that if X = total faces on n dice, $\sigma_x^2 = 0$,

when n is very large. But $\sigma_x^2 = n\sigma_d^2$, where σ_d^2 is the variance of one die. Hence as n gets large σ_x^2 becomes large and Spearman's proposition is false" (p. 332). This statement is a good illustration of how easy it is to demolish what one does not understand. Apparently, Tryon's difficulty lies in his failure to see that an individual's total ability remains constant whether one assumes 2 or 2,000 factors. That is, the range of ability within the group is not increased no matter how many factors are postulated. Each determiner or factor has an equal chance of being present (p) or absent (q), i.e., $p + q = 1$. The number of factors possessed by the average person within a given group is always 50 per cent, ($M = p = .5$); and the $\sigma = .5/\sqrt{n}$. Hence as n , the number of factors, increases, σ approaches 0, which means that more and more persons move in toward the mean, i.e., tend to possess 50 per cent of the total number of factors posited.

An illustration will show more clearly, perhaps, the truth of Spearman's conclusion. Suppose that in a very simple species, height is dependent upon one D (determiner); when D is present the individual is T (tall), when D is absent the individual is S (short). Now if our single D is split up into two smaller D 's, a third class of intermediate individuals will be added to our T 's and S 's. But it is highly improbable that we shall find individuals taller than our T 's or shorter than our S 's; namely, that the range will be increased. Otherwise we could 'increase' ability simply by increasing the number of postulated factors! Whether D is split up into 5 or 5,000 parts the only effect of increasing the number of factors is to increase the intermediate gradations between T and S . When there are 3 factors the probability of an extreme value (T or S) is .25, of an intermediate value .50; when there are 10 factors the probability of an extreme value is approximately .001, of an intermediate value .998. As the number of factors increases, therefore, the probability of intermediate scores becomes greater and greater and the probability of extreme scores less and less. This is, of course exactly what Spearman meant.

The operation of Spearman's principle may be observed

quite readily when one is studying the reliability of a given test. The more unreliable the test (the greater the chance elements entering into a score), the larger the standard error of a single score, and the more obscured the individual differences. When the standard error of a single score equals the standard deviation of the test (reliability coefficient equals 0), the most probable score for every person is the mean.

3. THE METHOD OF LOCATING GROUP FACTORS

The method whereby tests, which by the tetrad criterion appear to contain group factors over and above any general factor, are combined or dropped out in order to facilitate further analysis is called by Tryon (9) the 'purification technique.' His objection to this method seems to be that it (1) is *a posteriori* and (2) leads to results which are psychologically inconsistent. I do not think that either of these criticisms is valid. In many, perhaps all, psychological problems involving data taken from large groups, *a posteriori* methods must be employed first. When the psychologist attacks a new problem he may have to gather his data and then see what he can make out of them. But after this exploratory survey, he can usually select his tests or other instruments with a definite hypothesis in mind. This follow-up procedure is characteristic of Spearman's work, and is certainly true of the Columbia studies. Any scheme for minimizing small and ostensibly unimportant linkages is legitimate scientific procedure, if by so doing one is enabled the better to investigate more important connections which appear to be operative. An analysis¹ of Anastasi's memory data by Thurstone's center of gravity method has substantiated in clear-cut fashion the analysis made by the tetrad-difference technique. In fact, a comparison of Spearman's methods with those of Hotelling and Thurstone, when made upon the same data, indicates that the former instead of being "singularly crude," as Tryon states, are exceedingly precise.

¹ To be published shortly.

4. THE TWO-FACTOR VS. THE MULTIPLE FACTOR EXPLANATION OF INDIVIDUAL DIFFERENCES

In a strict sense it is incorrect to speak of the Two-Factor theory as though it were opposed to multiple factor theories. Since Spearman admits the existence of, and has systematically investigated, group factors with differing amounts of spread, his theory is truly a multiple factor theory. The Two-Factor theory does differ radically from other theories, however, in attempting to give a more parsimonious description of behavior. Spearman holds that the general factor (or *g*) far outweighs group factors in importance. Hence individual differences in performance are ordinarily attributed by him to *g* plus the factors specific to the given tasks (the *s*'s), rather than to a multiplicity of semi-general factors.

In place of the Two-Factor theory, Tryon, following Godfrey Thomson, proposes a multiple factor theory, namely, "that abilities are determined by a large number of more or less independent factors." In defense of this view, he advances the opinion that "these factors on the hereditary side are *genes*, and on the environmental side, they are the innumerable conditionings or associations formed in the course of learning" (8, p. 328). I do not believe that this interpretation of the causes of individual differences is as probable theoretically nor as well substantiated experimentally, as the explanation in terms of a few broad group factors or in terms of Spearman's *g* and *s*. It is undoubtedly true that ability-differences depend basically upon hereditary constitution. But the relationship of any given performance to its hereditary determiners is certainly unknown; and it is extremely unlikely that the relationship of mathematically determined factors to the *genes* is one of identity.

Neither theory nor observation favors the view that a learned act is the end-result of 'innumerable conditionings or associations.' Except perhaps at the very beginning of learning, associations are rarely specific S-R bonds. The doctrine of 'cue reduction,' the 'law of combination,' and the *Gestalt* description of learning all emphasize the idea of 'whole' reactions to 'partial' stimuli. The appearance in learning ex-

periments of such phenomena as 'short circuiting' and 'higher units' is evidence of the gradual emergence of large common bonds and the suppression of small specific ones. The term integration, so much used with reference to personality, clearly implies a common basis for a wide variety of activities. If behavior were a bundle of specific responses, the prediction of future acts from present performance would be clearly impossible.

School achievement offers a concrete illustration of the unity of much learned behavior. Though the correlation is not perfect, a boy's record in high school provides a good prediction of his probable achievement in college; achievement in arithmetic gives the best prognosis of achievement in algebra, etc. It is perhaps the large common language component in school work and in intelligence tests which explains why the I.Q. is so useful practically in educational prediction.

In addition to his own 'genetic' argument for multiple factors, Tryon (8) brings forward Garnett's mathematical proof (4) that even when a correlational hierarchy is established, it is possible to reproduce the intercorrelations without resort to a general factor; and Thomson's empirical demonstration (2) of this fact by means of cards and dice. Nothing further need be said about this evidence except that it is not new, and may be interpreted in favor of two factors as convincingly as in favor of multiple factors. In an exchange of views over twenty years, Spearman and Thomson have presented their respective arguments rather fully. These the reader may easily examine without resort to Tryon's interpretation. Most workers in the field of mental organization will agree, I believe, that the final truth of any factor theory can never be established upon purely mathematical or statistical grounds.

5. FACTORS VS. FACULTIES

Confusion has often arisen, I think, through the attempt of some psychologists to identify the factors obtained by statistical analysis of a test battery with the time-worn 'mental faculties.' Tryon is not the only offender in this respect. In a recent review of Perry's monograph, G. W. Allport (1) has

urged that an analysis of personality measures into factors is untenable, since personality traits are obviously not independent variables. Perhaps it would be well to look first at the classical definition of a faculty before trying to make clear what is meant by a statistically determined factor. Michael Maher, S.J., (5) defines a 'faculty' or 'power' as follows: "By a faculty is meant the mind's capability of undergoing a particular kind of activity; thus our sensations of colour are due to the faculty of vision, our judgments to the faculty of intellect, and our volitions to the faculty of will." It seems clear that a faculty is to be thought of, therefore, as a dynamic entity which controls or governs certain categories of behavior. 'Factors,' on the other hand, are mathematically determined constants which depend directly upon correlation coefficients, and represent the strength with which certain independent connections operate through the tests of a battery. This use of the term factor corresponds to the mathematical usage in which a common factor to x , y , and z is a quantity which enters (usually with unequal weights) into all three variables. When a 'verbal factor' is postulated in a battery of tests, a person's scores on these tests (apart from such influences as emotional stress, fatigue, ill-health, etc.) are considered to depend in part upon the degree to which he possesses this 'trait.' Conceived psychologically, factors represent the underlying unities within measured samples of behavior. Ultimately such unities must depend upon similar physical, physiological, and neurological structure.

A person's abilities may be described in a variety of ways. The physiologist, the psychiatrist, the social psychologist—each describes the individual from his own point of view. In like manner, the two-factorist describes the individual with respect to two dimensions g and s ; while the multiple factorist may select as many axes of reference as there are tests in his battery. Factor analysis offers a precise method of description with reference to a mathematically defined frame of reference. Which of the many possible descriptions of an individual is the 'best' one will depend, it would seem, upon the purposes of the investigator. Description in terms of

factors does possess, however, certain advantages in precision over purely verbal accounts. And it offers, too, decided advantages in that it affords a quantitative basis for prediction.

In summarizing this discussion, I think it may be fairly said that Tryon's criticisms offer no new or valid evidence against the Two-Factor hypothesis. At their best they are provocative; at their worst they illustrate highly successful attacks upon straw men. Probably no psychological theory has suffered more from misunderstanding or from partial understanding than has Spearman's. Uninformed critics have not hesitated to make dogmatic and sweeping statements regarding it. While Spearman has generally welcomed relevant criticism, it must become increasingly wearisome to have to fight the same battles over and over again. Under the circumstances he has displayed considerable forbearance.

REFERENCES

1. ALLPORT, G. W., Review of R. C. Perry's 'A group factor analysis of the adjustment questionnaire,' Southern California Education-Monographs, 1934, No. 5, pp. 93, in *Character and Personality*, 1934, 3, 169-170.
2. BROWN, W., AND THOMSON, G. H., The essentials of mental measurement, Camb. Univ. Press, 1921, pp. viii + 216 (Chapter X).
3. FLANAGAN, J. C., Factor analysis in the study of personality, Stanford Univ. Press, 1935, pp. x + 103.
4. GARNETT, J. C. M., The single *g* factor in dissimilar mental measurements, *Brit. J. Psychol.*, 1920, 10, 242-258.
5. MAHER, MICHAEL, Psychology—empirical and rational, Longmans, 1921, pp. xv + 603 + xii.
6. SPEARMAN, C., The sub-structure of the mind, *Brit. J. Psychol.*, 1927-28, 18, 249-261.
7. SPEARMAN, C., Professor Tryon on factors, *PSYCHOL. REV.*, 1934, 41, 306-7.
8. TRYON, R. C., Multiple factors *vs.* two factors as determiners of abilities, *PSYCHOL. REV.*, 1932, 39, 324-351.
9. TRYON, R. C., So-called group factors as determiners of abilities, *PSYCHOL. REV.*, 1932, 39, 403-439.
10. TRYON, R. C., Interpretation of Professor Spearman's comments, *PSYCHOL. REV.*, 1935, 42, 122-125.

[MS. received February 25, 1935]

PSYCHOLOGICAL REVIEW PUBLICATIONS

Original contributions and discussions intended for the *Psychological Review* should be addressed to

Professor Herbert S. Langfeld, Editor *PSYCHOLOGICAL REVIEW*,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the *Journal of Experimental Psychology* should be addressed to

Professor Samuel W. Fernberger, Editor *JOURNAL OF EXPERIMENTAL PSYCHOLOGY*,
University of Pennsylvania, Philadelphia, Pa.

Contributions intended for the *Psychological Monographs* should be addressed to

Professor Joseph Peterson, Editor *PSYCHOLOGICAL MONOGRAPHS*,
George Peabody College for Teachers, Nashville, Tenn.

Reviews of books and articles intended for the *Psychological Bulletin*, announcements and notes of current interest, and *books offered for review* should be sent to

Professor John A. McGeoch, Editor *PSYCHOLOGICAL BULLETIN*,
University of Missouri, Columbia, Mo.

Titles and reprints intended for the *Psychological Index* should be sent to

Professor Walter S. Hunter, Editor *PSYCHOLOGICAL INDEX*,
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company, Princeton, New Jersey

THE PSYCHOLOGICAL REVIEW

is indexed in the

International Index to Periodicals

to be found in most public and
college libraries

AMERICAN PSYCHOLOGICAL PERIODICALS

- American Journal of Psychology**—Ithaca, N. Y.; Cornell University.
Subscription \$6.50. 624 pages annually. Edited by M. F. Washburn, Madison
Bentley, K. M. Dallenbach, and E. G. Boring.
Quarterly. General and experimental psychology. Founded 1887.
- Journal of Genetic Psychology**—Worcester, Mass.; Clark University Press.
Subscription \$14.00 per yr.; \$7.00 per vol. 1,000 pages ann. (2 vols.). Edited by Carl
Murchison.
Quarterly. Child behavior, animal behavior, comparative psychology. Founded 1891.
- Psychological Review**—Princeton, N. J.; Psychological Review Company.
Subscription \$5.50. 540 pages annually. Edited by Herbert S. Langfeld.
Bi-monthly. General psychology. Founded 1894.
- Psychological Monographs**—Princeton, N. J.; Psychological Review Company.
Subscription \$6.00 per vol. 500 pages. Edited by Joseph Peterson.
Without fixed dates, each issue one or more researches. Founded 1895.
- Psychological Index**—Princeton, N. J.; Psychological Review Company.
Subscription \$4.00. 400-500 pages. Edited by Walter S. Hunter and R. R. Willoughby.
An annual bibliography of psychological literature. Founded 1895.
- Psychological Bulletin**—Princeton, N. J.; Psychological Review Company.
Subscription \$6.00. 720 pages annually. Edited by John A. McGeoch.
Monthly (10 numbers). Psychological literature. Founded 1904.
- Archives of Psychology**—New York, N. Y.; Columbia University.
Subscription \$6.00. 500 pages per volume. Edited by R. S. Woodworth.
Without fixed dates, each number a single experimental study. Founded 1906.
- Journal of Abnormal and Social Psychology**—Enc. Hall, Princeton, N. J.; American Psycho-
logical Association.
Subscription \$5.00. 448 pages annually. Edited by Henry T. Moore.
Quarterly. Abnormal and social. Founded 1906.
- Psychological Clinic**—Philadelphia, Pa.; Psychological Clinic Press.
Subscription \$3.00. 288 pages. Edited by Lightner Witmer.
Without fixed dates (Quarterly). Orthogenics, psychology, hygiene. Founded 1907.
- Journal of Educational Psychology**—Baltimore; Warwick & York.
Subscription \$6.00. 720 pages. Monthly except June to August.
Edited by J. W. Dunlap, P. M. Symonds and H. E. Jones. Founded 1910.
- Psychoanalytic Review**—Washington, D. C.; 3617 10th St., N. W.
Subscription \$6.00. 500 pages annually. Edited by W. A. White and S. E. Jelliffe.
Quarterly. Psychoanalysis. Founded 1913.
- Journal of Experimental Psychology**—Princeton, N. J.; Psychological Review Company.
Subscription \$7.00. 900 pages annually. Edited by Samuel W. Fernberger.
Bi-monthly. Experimental psychology. Founded 1916.
- Journal of Applied Psychology**—Indianapolis; C. E. Pauley & Co.
Subscription \$5.00. 600 pages annually. Edited by James P. Porter, Ohio University,
Athens, Ohio. Bi-monthly. Founded 1917.
- Journal of Comparative Psychology**—Baltimore, Md.; Williams & Wilkins Company.
Subscription \$5.00 per volume of 480 pages. Ed. by Knight Dunlap and Robert
M. Yerkes. Two volumes a year. Founded 1921.
- Comparative Psychology Monographs**—Baltimore, Md.; The Johns Hopkins Press.
Subscription \$5.00. 400 pages per volume. Knight Dunlap, Managing Editor.
Published without fixed dates, each number a single research. Founded 1922.
- Genetic Psychology Monographs**—Worcester, Mass.; Clark University Press.
Subscription \$7.00 per vol. One volume per year. Edited by Carl Murchison.
Bi-monthly. Each number one complete research. Child behavior, animal behavior,
and comparative psychology. Founded 1923.
- Psychological Abstracts**—Enc. Hall, Princeton, N. J.; American Psychological Association.
Subscription \$6.00. 700 pages ann. Edited by Walter S. Hunter and R. R. Willoughby.
Monthly. Abstracts of psychological literature. Founded 1927.
- Journal of General Psychology**—Worcester, Mass.; Clark University Press.
Subscription \$14.00 per yr.; \$7.00 per vol. 1,000 pages ann. (2 vols.). Edited by
Carl Murchison.
Quarterly. Experimental, theoretical, clinical, historical psychology. Founded 1927.
- Journal of Social Psychology**—Worcester, Mass.; Clark University Press.
Subscription \$7.00. 500 pages annually. Ed. by John Dewey and Carl Murchison.
Quarterly. Political, racial, and differential psychology. Founded 1929.

